

# **Aristotelian Causal Pluralism and Mental Causation**

Marc William Cole

Submitted in accordance with the requirements for the degree of Doctor of Philosophy.

The University of Leeds

School of Philosophy, Religion and History of Science

March 2019

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Marc William Cole to be identified as Author of this work has been asserted by Marc William Cole in accordance with the Copyright, Designs and Patents Act 1988.

### Acknowledgements

There are so many people to thank! I would like to start by thanking my wife Oksana and my children--Calvin, Svyatoslav, and Tadeusz. You all were critical in supporting me through this time. Indeed, I couldn't have done it without you.

To my supervisors, Heather Logue and Jamie Dow. Thank you for all you did to help me bring my idea to fruition, for all the meetings, for all the discussions.

I would also like to thank Helen Steward and Anna Marmodoro for examining my thesis. Thank you for all your helpful and thoughtful comments about my project! It was an honor to work with you.

To all my philosophy friends at the University of Leeds and beyond. Your friendship and support helped sustain me through dark times. Thank you, as well, for the countless philosophical discussions!

I would also like to thank the St Barnabas Society for stepping in and helping us at a critical juncture.

Thank you to my employer for all the support and understanding as I worked full time and finished my dissertation.

To Morgan with love.

*J. M. J.*

### Abstract

In my thesis, I argue that Aristotle's causal pluralism in conjunction with his hylomorphic account of mental states solves the exclusion problems for mental causation.

The exclusion problems are the result of a causal tension between the principle of Causal Closure (the thesis that caused physical events have sufficient physical causes) and the principle of Anti-reductionism (the thesis that mental properties and physical properties are distinct). The first exclusion problem is that Causal Closure excludes any non-physical causal influences on physical effects. The second exclusion problem shows that if one also holds any version of supervenience theory between mental and physical properties, mental states cannot cause other mental states. Denying either Causal Closure or Anti-reductionism would help to solve the problem.

Many consider Causal Closure in a stronger position than Anti-reductionism. However, I argue that there is good reason to deny Causal Closure and retain Anti-reductionism. This is only the first step. This just means that other, non-physical causes can influence the physical domain. But it says nothing about what these causes are or how they are related to their effects and perhaps other causes.

I suggest that a way forward is through causal pluralism, but, after putting forward a version of causal pluralism, I show it is still unclear how causes are related to other causes; in particular, how a mental cause is related to its underlying neural state.

I argue that Aristotle's hylomorphism helps with understanding how a mental cause is related to its underlying base, and his causal pluralism helps with how to understand a plurality of causes in mental causation. Aristotle's metaphysics of causation show how each of the causes are related to one another and how they are related to their effects. These insights are applied to mental causation and the exclusion problems.

## Table of Contents

<i><b>Introduction: Thesis, Context, and Project Outline</b></i>	<b>1</b>
<i><b>Chapter 1: The Exclusion Problems for Mental Causation</b></i>	<b>4</b>
<b>1.1 The Problem of Mental-to-Physical Causation</b>	<b>4</b>
<b>1.2 The Problem of Mental-to-Mental Causation</b>	<b>7</b>
<b>1.3 Probing the Principles</b>	<b>10</b>
1.3.1 Anti-reductionism and Causal Exclusion	10
1.3.2 Causal Closure	12
<i><b>Chapter 2: A Solution to the Exclusion Problem: Deny Causal Closure, Embrace Causal Pluralism</b></i>	<b>18</b>
<b>2.1 Causal Closure is Under-Motivated</b>	<b>18</b>
2.1.1 Causal Closure at the level of Physics	18
2.1.2 Causal Closure in the Special Sciences	27
<b>2.2 A Causal Pluralist Alternative</b>	<b>30</b>
2.2.1 Causal Pluralism in Broad Outline	30
2.2.2 Causal Pluralism and Causal Closure	34
2.2.3 Causal pluralism and Causal Exclusion	36
2.2.4 Rationalising vs. Triggering Causes	38
<b>2.3 Lingering Worries</b>	<b>39</b>
2.3.1 How Are the Rationalizing and Triggering Causes Related to Each Other?	40
2.3.2 Why Don't These Two Kinds of Causes Compete for Sufficiency?	44
2.3.3 Jaworski's Neutrality on Causal Closure	47
<b>2.4 Summing Up and Some Questions for Jaworski</b>	<b>49</b>
<i><b>Chapter 3: Aristotle's causal pluralism</b></i>	<b>51</b>
<b>3.1 The Relation between Explanation and Causation on Aristotle's View</b>	<b>51</b>
<b>3.2 Aristotle's Four Causes</b>	<b>57</b>
3.2.1 The Material Cause	58
3.2.2 The Formal Cause	60
3.2.3 The Efficient Cause	66
3.2.4 The Final Cause--The "End" of the "Matter"	69
3.2.5 The four causes at work in explaining change	70
3.2.5.1 Material Cause in Change	71

3.2.5.2 Formal Cause in Change	72
3.2.5.3 The Efficient Cause in Change	74
3.2.5.4 The Final Cause in Change	76
3.2.6 Taking Stock, and Some Loose Ends	77
<b>3.3 Why the Four Causes do not Compete</b>	<b>79</b>
<b>3.4 How Matter and Form are Related</b>	<b>83</b>
<b><i>Chapter 4: How Aristotle's Causal Pluralism Solves the Exclusion Problems for Mental Causation</i></b>	<b>89</b>
<b>4.1 Aristotle's Four Causes and Mental Causation</b>	<b>90</b>
4.1.1 Aristotle's Account of Psychological States	90
4.1.2 The Four Causes at Work in a Case of Mental Causation	99
4.1.2.1. Mental causation without deliberation	100
4.1.2.2 Mental causation with deliberation	103
4.1.2.3 Some Further Thoughts on Aristotle and Desire in Action	112
4.1.2.4 Mental-to-Mental Causation	113
4.1.3 Why the Causes Do Not Compete in Cases of Mental Causation	114
4.1.3.1 Why the Causes are not in Competition for Mental-to-Mental Causation	124
4.1.4 How Form and Matter are Related in the Case of Mental States	124
<b>4.2 Mapping Jaworski's Rationalizing and Triggering Causes into the Aristotelian Framework</b>	<b>136</b>
4.2.1 Placement of the Rationalizing Cause in Aristotle's Metaphysics	137
4.2.2 Placement of the Triggering Cause in Aristotle's Framework	143
4.2.3 How the Rationalizing and Triggering Causes are Related, and Why They Are Not in Competition	144
<b><i>Chapter 5: Conclusion</i></b>	<b>148</b>
<b><i>References</i></b>	<b>153</b>

## Introduction: Thesis, Context, and Project Outline

The thesis of this project is that Aristotle's causal pluralism, including hishylomorphic account of mental states, solves the exclusion problems for mental causation and offers an account of how mental properties are related to their neural bases and their subsequent effects.

The context of this project takes its cue from the exclusion problems for mental causation, especially as formulated by Jaegwon Kim. The exclusion problems demonstrate the problems for "finding a place" for mental properties to be causally efficacious in the world (a desire for health causes walking, a desire to entertain causes singing, etc.). The shape of the problems can be set up in the following way. There are plenty of reasons for thinking that mental properties are distinct from physical properties and the thesis that upholds this can be called Anti-reductionism (following Kim's terminology). However, given the seemingly highly plausible thesis of the Causal Closure of the physical domain, it would appear that mental properties are *excluded* from the causal picture on the grounds that physical events, if they have causes, have sufficient physical causes. If this right, then physical events (walking, singing, etc.) that seem to be caused by beliefs or desires etc, are really caused by the physical neural states that underlie mental properties.

Finding a theory that allows mental states to be distinct from physical states and also causally efficacious while holding Causal Closure has proved to be very difficult. So, there is an incentive to try to solve the problem by denying either Causal Closure or Anti-reductionism. Denying Anti-reductionism results in a theory where mental states just are physical states (as the identity theorists would have it) or that mental states and many of our folk psychological concepts are just illusions to be replaced by a future-completed neuroscience (as the eliminative materialists would have it). Either way of denying Anti-reductionism solves the problem. Identity theorists, by identifying mental

states with physical states, can directly explain how mental states are causally efficacious: they are causally efficacious in virtue of being physical. Eliminative materialists solve the problem by saying there aren't any folk psychological states that are causally efficacious in the first place. Such moves are highly controversial, however—it is far from clear that mental properties are the sorts of things that can be identified with physical properties, and it is difficult to accept the idea that folk psychological concepts such as beliefs and desires are merely illusory.

So we have at least some reason to seriously consider solving the problem by denying Causal Closure. By allowing the possibility that the physical domain can be causally influenced by non-physical causes, there is room made for mental properties to be causally efficacious in bringing about physical states. But many find this a difficult solution to accept as well. For one thing, Causal Closure has a venerable history in science, and there appears to be good evidence for it. But even if this obstacle could be surmounted, there are still lingering questions of detail regarding how mental states would be related to their neural bases on the one hand, and exactly what role they play in bringing about their effects on the other.

In this project, I want to pursue the path of denying Causal Closure. In particular, I will explicate a plausible interpretation of Aristotle's theory of causal pluralism andhylomorphism, and show how it enables us to solve the exclusion problem by denying Causal Closure, while also giving us the resources to explain how mental states are related to neural activity and what role they play in bringing about their effects.

This project proceeds in the following way. First, I introduce the exclusion problem for mental causation as presented by Jaegwon Kim. The exclusion problem has two aspects: mental-to-physical causation (how can mental states cause physical states?) and mental-to-mental causation (how can mental states cause other mental states?). In



Kim's version, we should make a choice between Anti-reductionism and Causal Closure because holding both does not seem possible. He thinks that it is obvious that Causal Closure has the better claim, so we should lose Anti-reductionism. I argue that Anti-reductionism is actually better motivated than Causal Closure and that much of the power of Causal Closure seems to be a hangover from an outdated view of causation (“causation-as-production”).

In chapter 2, I argue that there is actually no good reason to accept Causal Closure based on the works of Helen Steward and Nancy Cartwright. Following this, I introduce William Jaworski's causal pluralism as a potential alternative, on which neural states are “triggering” causes and mental states are “rationalising” causes. However, Jaworski is neutral on whether or not Causal Closure holds, and this leaves him with two difficult questions: (1) how are rationalizing causes related to triggering causes, and (2) why aren't rationalizing and triggering causes in causal competition?

In the remainder of the thesis, I argue that these questions can be answered if we adopt Aristotle's causal pluralism andhylomorphic account of mental states. In chapter 3, I articulate these views in detail. In chapter 4, I show how Aristotle applies these views to mental states and explain how this helps answer the questions of how rationalising and triggering causes are related to each other, and why they aren't in causal competition. The resulting Aristotelian view allows us to solve the exclusion problem for mental causation by denying Causal Closure.

## Chapter 1: The Exclusion Problems for Mental Causation

### 1.1 The Problem of Mental-to-Physical Causation

Morgan is crying. If you were to ask her why, she would say she is crying because of philosophical angst. More specifically, Morgan would say she is crying because she wants to say mental states are causal; but, in Philosophy of Mind class today, she learned that given some very reasonable assumptions about causation, it would seem mental states are excluded from the causal picture. So, she wants to say she is crying because of this angst, but good causal assumptions seem to exclude precisely this. In short, Morgan is crying about the exclusion problem. What are these “reasonable assumptions about causation?” Jaegwon Kim (2011) gives them here, along with the mental/physical distinction:

- (1) “*Causal Closure of the physical domain*. If a physical event has a cause at  $t$ , it has a sufficient physical cause at  $t$ ” (Kim, 2011, p. 38).
- (2) “*Anti-reductionism*. Mental properties are distinct from, and irreducible to, physical properties” (Kim, 2011, p. 39).
- (3) “*Causal Exclusion*. No event can have more than one sufficient cause occurring at any given time—unless it is a genuine case of causal overdetermination” (Kim, 2011, p. 39).

(2) is straightforward; it just means that mental and physical properties are distinct. This is what Morgan pretheoretically accepted before class today. Some further considerations are in order. When Kim says that mental properties are irreducible to physical properties, I take it that he means that mental properties cannot be explained by the resources of physical properties. I would think it does not mean that there is no connection between mental and physical properties. Perhaps physical properties somehow give rise to mental properties; or perhaps mental properties supervene on

physical properties. (2) can accommodate such views. It would just mean that mental properties are not fully reducible to physical properties. At any rate (2) is flexible enough to accommodate weaker or stronger versions of (2).

(1) is the first reasonable causal principle Morgan learned about. It means that if a physical event, like Morgan crying, has a cause at all, it will have a sufficient physical cause. (3) is the other reasonable causal principle Morgan learned about today. It means that only one sufficient cause occurs at a given time, unless there is a genuine case of overdetermination wherein more than one sufficient cause is in effect. So, in our example, if Morgan's crying has a sufficient cause, then it only has *one* sufficient cause, unless there is a case of overdetermination.

These three premises present us with an inconsistent triad. Morgan wants to hold (2); she wants to say her angst is distinct from, and irreducible to, her brain state. But she also wants to hold that this mental state of angst, M, caused her to cry. For the moment, let's focus on this latter claim: M caused her to cry. Also, let the underlying brain state of M be called  $p_1$ , where "p" is for "physical." And let the physical act of crying be called  $p_2$ . Morgan wants to say that M caused  $p_2$ . She wants this to be the case for her particular instance of angst, but she also wants the schema to generalize to any mental state that seems to cause a subsequent physical effect. For example, my hunger caused me to eat; my decision caused me to enter grad school, my love caused me to hug my son.

The problem is that she also thinks (1) and (3) are good and reasonable assumptions about causation. (1) says that if  $p_2$  has a cause, then it has a sufficient physical cause. Well, the only likely candidate in the neighborhood is  $p_1$ , the underlying neural state of M. Morgan could try to say that M is also a sufficient cause, thus making  $p_2$  overdetermined. But this seems like a bad result. This would mean that certain mental properties, such as my decision to enter grad school, my love, or my hunger are

sufficient causes of filling out an application, reaching out to hug my son, and getting up to go to the refrigerator respectively. But the underlying physical states are also sufficient. The upshot is that a systematic overdetermination of  $p_2$  is required to make this work. But this seems like a bad result, for it seems terribly *ad hoc*. The only reason one would posit such overdetermination is to keep (2). But (1) gives us good reason to strongly suspect (2). A further argument for systematic overdetermination is therefore needed. And so, (3) kicks in here. Without a good reason to suppose systematic overdetermination occurs, which no one including Morgan, really wants, then we have to choose between keeping the mental cause or physical cause of  $p_2$ . But as Kim (2011) observes, “Excluding the physical cause will not do any good, since if we try to retain the mental cause, physical Causal Closure will kick in again and bring back the excluded physical cause” (p. 39). This is because  $p_2$  is a physical event, and, according to (1), any physical event, if it has a cause, has a sufficient physical cause. The only way to exclude the physical cause would be to deny (1). And so, we have to let mental causes go; they are excluded. Kim (2011) summarizes thus:

“Physical Causal Closure implies that any putative non-physical cause of a physical event will face competition from a physical event for causal status. The exclusion principle comes into play at this point, yielding the conclusion that in such competition the physical cause always prevails” (p. 39).

And so mental-to-physical causation seems ruled out by Causal Closure.

There are other problems for many Anti-reductionists. In broad strokes, there are at least two ways to be an Anti-reductionist. One could be a property dualist. The standard understanding of property dualism is that mental properties are non-physical properties (although most property dualists would say mental properties are emergent properties). These views usually bring with them a variety of philosophical baggage:

how are non-physical mental properties related to a physical property, such as a neural state? How can a non-physical property cause a physical event ( $p_2$ )?<sup>1</sup> But the standard way of being an Anti-reductionist is by being a non-reductive physicalist.

## 1.2 The Problem of Mental-to-Mental Causation

Unlike most kinds of property dualism, Non-reductive physicalism argues that mental properties depend in some way on physical properties.<sup>2</sup> Usually, this dependence relation is discussed in terms of supervenience. At any rate, Non-reductive physicalism, according to Lynne Rudder Baker (2011), has at least three theses:

“(1) There are mental properties that are distinct from any physical properties.

(2) Mental properties depend on physical properties.

(3) Mental properties make a causal contribution to what happens” (pp. 110-111).<sup>3</sup>

As Baker (2011) rightly notes, this is only a schema: there are a variety of ways to understand “make a causal contribution to” in (3); for example, nomological sufficiency or counterfactual dependence (p. 111). Likewise, says Baker (2011) the “depend on” relation in (2) can mean “weakly supervene on,” “globally supervene on,” or “strongly supervene on,” to name some of the more popular views (p. 111). Indeed, some version of supervenience is the orthodoxy in accounting for the dependence of the mental on the physical. It is this point of supervenience that Kim targets. But first, it merits discussion to look at the inconsistent triad of the last section in light of Non-reductive physicalism.

---

<sup>1</sup> Property dualism is the successor of Rene Descartes’ substance dualism. His substance dualism posited that the mind and body were of two different substances: the body was a physical substance hallmarked by extension, and the mind was of another substance hallmarked by thinking and not extension. For an interesting discussion of the problems besetting Cartesian dualism, see: Kim (2011, pp. 30-35). Many of the issues there translate over to property dualists, especially on the questions mentioned in the text.

<sup>2</sup> The brilliant forerunner of contemporary Non-reductive physicalism is Donald Davidson’s (2002) publication of “Mental Events,” in which he introduces anomalous monism (originally published in 1970).

<sup>3</sup> Baker calls her theory “Non-reductive materialism.” I am using the term “Non-reductive physicalism.” I do not think much hangs on the terminology for my purposes here.

For convenience, here again is the inconsistent triad of the previous section:

- (1) “*Causal Closure of the physical domain*. If a physical event has a cause at  $t$ , it has a sufficient physical cause at  $t$ ” (Kim, 2011, p. 38).
- (2) “*Anti-reductionism*. Mental properties are distinct from, and irreducible to, physical properties” (Kim, 2011, p. 39).
- (3) “*Causal Exclusion*. No event can have more than one sufficient cause occurring at any given time—unless it is a genuine case of causal overdetermination” (Kim, 2011, p. 39).

Even if Non-reductive physicalism differs from property dualism by positing that mental properties depend on physical properties, Non-reductive physicalism still holds (2) Anti-reductionism. It is specifically this thesis that introduces problems when held together with (1) and (3). As long as a mental property has to compete with a physical property for being the sufficient cause for a subsequent physical event, exclusion looms dark on the horizon. Anyone who starts out holding (1)-(3) will have to make some hard causal choices, namely excluding mental causes given the seeming power of Causal Closure.

But Kim has another argument for those who hold to some version of supervenience. The argument shows that mental-to-mental causation fails. Here is Kim’s (2011) version of the supervenience thesis:

- (4) “*Mind-Body Supervenience*. Whenever something has a mental property,  $M$ , at  $t$ , it does so in virtue of the fact that it has, at  $t$ , a physical base property,  $P$ , where  $P$  necessitates  $M$  (that is, necessarily anything that has  $P$  at  $t$  has  $M$  at  $t$ )” (Kim, 2011, p. 40).

Recall that supervenience is supposed to explain the dependence relation between a mental property and a physical base property. (4) says the reason there is a mental

property M in the first place is because of a physical base property P, such that anything with P at *t*, has M. To put it more crudely, if you have P at *t*, you must also have M at *t*.

The problem for the holder of supervenience, according to Kim (2011), is that supervenience implies downward causation (p. 40), which the exclusion argument shows is not possible. Suppose Morgan is feeling angst. Her angst (M) is caused by thinking about philosophy (M\*). But supervenience says that M occurs in virtue of a physical base, P, and M\* occurs in virtue of a suitable physical base, P\*. The only way for M\* to cause M would be in virtue of causing P.

Downward causation might not seem like a bad result to some, but when combined with the exclusion argument in the previous section, the holder of downward causation has a lot to answer for. Because of Mind-Body Supervenience, P\* (the neural state necessitating thinking philosophically) is sufficient for M (angst) in virtue of being sufficient for P (the neural state necessitating angst). Hence, the Causal Exclusion principle rules out M\* (thinking philosophically) as a cause of M (angst).

In conclusion, Morgan is crying. She wants to say that thinking about philosophy, M\*, caused her angst, M, which in turn, caused her to cry. But given some reasonable assumptions about causation, it seems that a physical base, P\*, caused P, which in turn caused the crying. This is the exclusion problem. And even if she took up a supervenience theory, it would not get her far. Supervenience shows that the only way a mental property could cause another mental property, M, would be to cause the underlying physical property, P. But then the exclusion problem demonstrates that mental-to-physical causation does not work anyway. At best, it looks like mental properties are epiphenomenal to their physical bases since it is the physical bases that are doing the causal heavy lifting. At worst, there are simply no mental properties at all, slaked off by the demands of Causal Exclusion and Causal Closure.

Poor, poor Morgan. What can be done? Next, I turn to some assumptions about causation on the one hand, and the mental/physical distinction on the other. After that, I introduce Aristotle's approach and, later, how it solves the exclusion problems. This should transform Morgan's melancholia into eudaimonia.

### 1.3 Probing the Principles

The problems of mental-to-physical and mental-to-mental causation have proved exceptionally difficult in philosophy of mind and there is no widely agreed-upon solution. The problem is sorting which of the theses should be denied. Should it be Anti-reductionism, Causal Closure or Causal Exclusion?

In what follows in this section, I want to motivate retaining Anti-reductionism and Causal Exclusion and argue that the motivation for Causal Closure is a hangover from a very narrow view of causation. In particular, a causation-as-production view, and an implied necessitation relation between cause and effect, are doing a lot of the motivational work in Causal Closure (and Causal Exclusion for that matter). However, causation-as-production and necessitation theories are far from the only views on the road these days and if some of these other views are plausible, the motivation for Causal Closure is undercut.

#### 1.3.1 Anti-reductionism and Causal Exclusion

Let us begin with Anti-reductionism. There are many arguments for it, but my aim here isn't to definitively defend any particular one of them, but rather to offer reasons why it is at least plausible. One good reason to hold Anti-reductionism is that mental properties really seem to be something fundamentally different than the physical materials and processes that constitute them.

Heil (2004) puts the point as follows:



“Imagine that you are attending a performance of Die Walkure at Bayreuth.

Your senses are assaulted by sounds, colors, smells, even tastes. A neuroscientist observing your brain while all this is occurring would observe a panoply of neural activities. But you can rest assured that the neuroscientist will not observe anything resembling the qualities of your conscious experience” (p. 2).

“Mental events, states, and properties appear to be utterly different in kind from material events, states, and properties. The difference is striking when we consider the qualities of material bodies, including the qualities of brains. Your visual experience of a ripe tomato in bright sunlight seems qualitatively very different from goings-on in your nervous system. Neurological occurrences can be observed and described in great detail. But observe as we will, we seem never to observe anything at all like a conscious experience” (p. 75).

The thought is that goings-on in the nervous system can be described in terms of electro-chemical processes occurring throughout the system. This can all be discovered and described through the usual scientific ways. Mental events seem to be different on at least two scores. First, mental events cannot be observed in the ways that neural events are (e.g., in an MRI machine). Second, electrochemical reactions seem like the wrong vehicle to explain the very particular feeling of seeing bright sunlight on tomatoes (and perhaps the feelings it kicks up; e.g. summer nostalgia, etc.) These considerations have led many, including Heil (2004), to hold that mental states are not built up out of physical entities. If the foregoing is right, then Anti-reductionism is quite plausible.

Let us now turn to the Causal Exclusion principle. In Kim's set-up of the exclusion problems, the thesis of Causal Exclusion is employed: an effect cannot have more than one sufficient cause unless it is a case of genuine overdetermination. Couldn't this thesis be rejected, and the exclusion problem solved? That is, and more to the point of this

project, could one reject the Causal Exclusion principle and solve the exclusion problem by accepting that cases of mental causation are just genuine cases of overdetermination? This solution requires, in effect, a rejection of Causal Closure. It would mean that if a mental cause somehow occurred by itself (without an underlying neural base) it would be sufficient to bring about physical action. So this proposed solution to the problem isn't really distinct from the option of denying Causal Closure. The real action is the choice between Causal Closure or Anti-reductionism. I turn now to the principle of Causal Closure.

### 1.3.2 Causal Closure

Causal Closure says that any physical event, if it has a cause, has a sufficient physical cause. It would be instructive to discuss what sorts of things count as physical.

The working definition of “physical” can be: Something counts as physical if it is an entity postulated by a physical theory (like an atom, or some more fundamental physical entity), or an entity built out of those entities (a living cell, for example).<sup>4</sup> This seems fairly straight-forward, but what of other domains that do not appear to be causally closed? David Papineau (2011) gives several examples of domains that are not causally closed:

“The biological realm, for example, is not causally closed...since biological effects often have non-biological causes, as when the impact of a meteorite precipitated the extinction of the dinosaurs. Again, meteorology is not causally closed: the burning of carbon fuels—a non-meteorological event—is causing global warming. Nor, importantly, is the mental realm causally closed: a mental pain can be caused by sitting on a physical drawing pin, or a train of thought can be interrupted by a loud noise” (p. 54).

---

<sup>4</sup> I would like to thank Heather Logue for help with this definition.

It seems to many that physics is causally closed in a way the other domains are not. But each of the domains cited have physical effects. Even though the burning of carbon fuels is not meteorological, it has a physical effect on the atmosphere. Even though sitting on a tack is not a mental event, it has the physical effects of jumping up and crying out. The meteorite, though non-biological, still had the physical effect of killing the dinosaurs. And this seems to be true even for physical effects that occur in living bodies. Suppose I arc my hand through the air. Papineau (2011) says:

“When the muscle fibres in my arm contract, this is presumably due to electrochemical activity in my nerves, which is due to prior physical activity in my motor cortex, and so on. In principle, it would seem possible to account for this entire sequence solely in terms of the resources offered by physics itself, and without making any essential appeal to any other subject matter” (p. 54).

So, on the one hand, the domain of physics appears to be causally closed. On the other hand, domains such as biology, meteorology, and psychology appear to be causally open. But each of these “open” domains appears to have effects on the physical domain. How can this be? The typical answer is that gases, tacks, and dinosaurs are all composed of entities postulated by physical theory.

For this reason, Causal Closure offers a good argument for reducing other domains to physics. Papineau (2011) says:

“...[T]he Causal Closure of the physical does give rise to a powerful argument for reducing many *prima facie* non-physical realms to physics: for it indicates that *anything that has a causal impact on the physical realm* must itself be physical. This is because the Causal Closure of the physical seems to leave no room for anything non-physical to make a causal difference to the physical realm, since it specifies that every physical effect already has a physical cause” (p. 54) (*italics Papineau’s*).

The thought is that, since physical effects must have sufficient physical causes, then those domains that cause physical effects, must themselves be physical. Papineau (2011) continues:

“An infection (biological) can cause a rise in temperature; a hurricane (meteorological) can destroy physical houses; my current thoughts (mental) can give rise to physical movements of my fingers on a keyboard. However, the Causal Closure of the physical says that these special effects must already have physical causes: the rise in temperature will be caused by cellular-level chemical processes; the houses’ destruction by fast moving gases; the movements of my fingers by neuronal activity in my brain. How then can the infection, the hurricane, and my thoughts also make a causal difference? The Causal Closure of the physical would seem to leave no room for these *prima facie* non-physical causes to matter to the physical effects” (pp. 54-55).

The critical point is that any of the other putative domains have physical causes and effects, and Causal Closure says that for any physical effect, there is a sufficient physical cause (if there is a cause at all). This strongly implies that the causes at entities of the special scientific domain are physical, or have a physical ontology, themselves. The infection’s effect can be explained in terms of cellular-level chemical processes<sup>5</sup> and the house was destroyed by fast-moving gases. Papineau (2011) agrees:

“...they [biological, meteorological, and mental causes] will be ‘non-physical’ only in the sense that they are normally referred to using specialist (biological, meteorological, mental) terminology, and not because they are ontologically different” (p. 55).

---

<sup>5</sup> Papineau seems to take it for granted that chemical processes are reducible to the physical domain. But this is debatable. See for example: Robin Findlay Hendry (2006). And chemistry is not the only example of a domain that might not be reducible to physics. More is said on this in chapter 2.

The ontology of the scientific special domains is just more of the same physical entities and causes. Papineau (2011) says that we can

“...identify the prima facie non-physical causes with the physical causes. If we equate the infection, hurricane, and thoughts with the cell-level activity, fast-moving gases, and neuronal activity respectively, then their causal efficacy will be respected. The biological, meteorological, and mental causes won't be eclipsed by the physical causes, simply because they will one and the same as the physical causes” (p. 55).

According to the argument at hand, then, what counts as physical are things and processes like the fast-moving gases of a hurricane, the cell-level activities, and the neuronal activities. These are comprised of physical entities and causally impact the physical domain. As such the domains of biology, meteorology, psychology, etc. are really just parts of the physical domain. Something counts as physical if it is an entity postulated by physics, or an entity built up out of the entities postulated by physics; and biological, meteorological, psychological (etc.) entities are built up out of those entities.

But of course, there is still the case for Anti-reductionism about the mental discussed in the previous section. Reductionism about the mental is not as plausible as the reduction about the meteorological, biological, etc.<sup>6</sup> So it would seem that if we're going to hang onto Anti-reductionism and Causal Exclusion, we'll have to give up on Causal Closure of the Physical.

Fortunately, there seems to be conceptual room to do that. I suspect that much of the motivation behind Causal Closure rests on an outdated view of causation, which we can term "causation-as-production."

Here again is the Causal Closure principle:

---

<sup>6</sup> And, it is not clear that even many of the special domains reduce to entities postulated by physicists. There will be a bit more on this in the next chapter.

"Causal Closure of the physical domain. If a physical event has a cause at  $t$ , it has a sufficient physical cause at  $t$ ."

The most straightforward way of reading Causal Closure is in a causation-as-production way. If a physical event,  $E$ , has a cause at all, it has a physical cause,  $P$ , such that  $P$  brings about  $E$ . "Brings about" here is most straightforwardly understood as "produces" in some way. Thomas Kroedel (2015) observes: "...the exclusion problem is most pressing for a view of causation as production, or causal 'oomph', which involves continuous spatio-temporal processes and the transfer of energy between cause and effect" (p. 369). A few examples can illustrate this. Imagine a factory that produces juice, imagine a watch, and imagine a carburetor in an internal combustion engine. The juice factory has a number of sequenced processes: from squashing the fruit, to filling bottles, and conveying the product from room to room. Each process is in direct contact with its effect. What squashes the fruit is presumably automated, following the commands issued to it. It also comes into contact with the fruit during the squashing process. And similarly for the other processes. Each of these are continuous spatio-temporal processes, and there is a transfer of energy between cause (the command to squash) and effect (squashing). A traditional watch has a number of cogs, rotors, and movements that are in continuous spatio-temporal contact, and therefore directly transfer energy among themselves in keeping time. Each part is connected to its effect. Finally, a carburetor mixes vaporized fuel and air in order to make a more combustible mixture for an internal combustion engine. The mixing is directly connected with its effect, the combustible mixture. What all these examples share in common are spatio-temporally connected processes and transfers of energy.<sup>7</sup>

---

<sup>7</sup> One might think that causation-as-production implies a kind of necessitation relation between events. In the examples I cited for example, the causes seem to determine their effect. But causation-as-production is compatible with probabilistic causal theories too. If we say "smoking causes cancer" where "causes" can be understood as "increases the risk of", there is still continuous spatio-temporal contact between the smoke of the cigarettes and the relevant body parts (such as mouth, lungs, and throat, say). The problem

Why does Kroedel say that the exclusion problem is most pressing for causation-as-production? Let's look at Morgan's underlying neural state of her angst,  $p_1$ , and the subsequent tears spilling,  $p_2$ . Taking an overly simplistic view, one might think that  $p_1$  and  $p_2$  are not unlike a carburetor or a watch. What occurs at  $p_1$  directly brings about  $p_2$ .<sup>8</sup> Now suppose there are two "competing" causes: the mental cause and the physical cause. The mental cause will always be excluded (assuming Causal Exclusion) because a physical cause is by far the most plausible candidate for a "continuous spatio-temporal process" and/or the source of an energy transfer. And it is most plausible because in many, many cases, such as the ones I cited, physical causes directly produce their effects.

But causation-as-production (or by contact) is not the only causal show on the road. Causal pluralists, for example, think there are many different kinds of causes (not just physical) and that these can be related to effects in many different ways (not just causation-as-production).<sup>9</sup> This will be discussed in greater detail in the next chapter, but for my purposes here it is enough to note that the kind of causal pluralism I am advocating recognizes other sorts of causes than those at work in causation-as-production and that we should deny Causal Closure of the physical domain. Denying Causal Closure is the subject of the next chapter.

---

with causation-as-production is that not all causes and their events seem related through production. This is discussed more in the causal pluralism section.

<sup>8</sup> The example is overly simplistic for a variety of reasons. For one thing, suppose Morgan was feeling philosophical angst, but did not want to cry for some other reason. And this other reason (presumably with its concomitant physical base) "overrides" the connection between  $p_1$  and  $p_2$ . For another thing, and relatedly, it is not clear if there is a clear-cut connection between  $p_1$  and  $p_2$  in the same way as the carburetor or watch. Nevertheless, it is still a sensible question to wonder if neural causes necessitate their effects, even if this turns out to be much more complicated than originally supposed.

<sup>9</sup> William Jaworski (2011) defines causal pluralism along these lines. Causal pluralism, says Jaworski (2011), is the thesis that "there are many different kinds of causes and many different kinds of causal relations" (p. 360).

## Chapter 2: A Solution to the Exclusion Problem: Deny Causal

### Closure, Embrace Causal Pluralism

In this chapter, I am going to argue that there is no compelling reason to accept Causal Closure of the physical domain, and then argue that some version of causal pluralism should be accepted. In particular, I introduce William Jaworski's view of causal pluralism. After discussing this, I show how he thinks that in instances of mental causation, there are rationalizing causes and triggering causes that correspond to mental and physical causes respectively. However, merely adopting causal pluralism as well as the rationalizing and triggering causes doesn't answer all the required questions that help to generate the exclusion problem.

#### 2.1 Causal Closure is Under-Motivated

The main contention of this section is that there is no particularly good reason to suppose that physics is causally closed. First, I will discuss an argument to the effect that not all events construed in uncontroversially physical terms have physical causes. Next, I will argue that the situation is analogous with respect to the special sciences.

##### 2.1.1 Causal Closure at the level of Physics

What is supposed to be so bad about denying Causal Closure? Kim (2011) says that denying Causal Closure

“...would imply that there might be non-physical causal agents, outside the physical domain, injecting causal influences into it, and that there might be physical events whose explanations must invoke these non-physical causal agents and forces” (p. 38).

The violation of physical laws via non-physical causal influence would have the apparently most unwelcome consequence that theoretical physics would, says Kim (2011), “...be in principle incomplete” (p. 38).



What would it mean if theoretical physics was indeed incomplete?

Presumably, it would mean that not every physical event would have a sufficient physical explanation. The thought is that the physical domain is closed to any non-physical kind of cause; anything that occurs within the physical domain can be explained employing the resources of physics and physics alone. So a "violation" of the physical domain means admitting of non-physical causes that are essential in explaining a caused physical event.

The "completeness" of physics has had an influential hold on modern science since arguably Descartes. It peaked with Newtonian physics and continues to hold sway well into the present day. However, there are good reasons to think that things are not so clear cut.

There are many contexts wherein the behavior of some physical stuff is influenced and/or determined by higher-order influences which do not result in a violation of physical laws. Helen Steward (2014) gives two examples: First, "...the physical influences on a molecule embedded in a cell, say, will be different from the physical influences on one that is not so embedded" (pp. 229-230).

The (biological) forces exerted on a molecule within a cell will cause that molecule to behave differently than if that molecule were sitting on its own apart from a living body. But this does not necessarily mean that the biological forces acting on the en-celled molecule are violating physical laws at the level of the molecule. Second, and similarly, Steward (2014) says that the molecules in a (frequently-used) door behave differently than they would if they were not part of a door (p. 230). That is, the molecules in the door will move backward and forward along very particular paths each time the door is opened and closed. The same molecules, if un-en-doored, will behave differently. But, as Steward (2014) rightly quips: "But surely no one thinks that any physical laws are violated when doors are opened" (p. 230).

What Kim and others would say in return is that biological laws and the forces involved in door-opening are well within the realm of physics, whereas mental states, as Anti-reductionists conceive them, are non-physical. The reason, then, that the molecule within the cell and the molecules in the door do not violate physical laws even though they are caught up in higher-order levels is because they are ontologically physical (*i.e.*, made up of entities postulated by physical theory).

Kim and others also have residual concerns about the nature of 'physical'. Steward (2014) thinks, and I agree, that Kim's concern would be better expressed in the following way:

“...Purely physical laws (*i.e.* laws expressible using only the concepts of *physics*) are sufficient to dictate completely the movements of all physical things. Cells and molecules are certainly physical things, and so purely physical laws ought to be sufficient to dictate their movements too. But then, unless psychological states and processes are themselves describable in physical ways, so that their influence on events, such as it is, can be found simply to be a form of influence that comes within the ambit of those physical laws, their having an impact on the way cells and molecules move is inconsistent with the assumption that everything is dictated by the physical laws. There will be extra, non-physical influences to take account of in the story of what causes motion in our universe. And that is something that many naturalistically inclined philosophers will be very disinclined to accept” (p. 230).

The reason that the cell and door examples above do not constitute the violation of physical laws is because doors and cells are ultimately made up out of physical entities, such as atoms and so on. So, the thought is that since doors and cells are physical, the physical laws (of physics) should ultimately be able to explain even

higher-order physical phenomena. This is the heart and soul of Causal Closure of the physical domain. The problem with mental states is they are often defined in non-physical terms and are supposed to have a causal impact on the physical domain according to exponents of Anti-reductionism. In the next section, I discuss that this only delays the problem for the advocate of Causal Closure; if the other domains are not reducible to physics, and their resources have an essential role to play in explaining what happens in some physical events, then Causal Closure still fails.

Elsewhere, Kim (1992) elaborates on the concern with mental states (p. 120).

Steward (2014) quotes and comments on the passage:

“The idea is that when certain wants and needs, aided by perceptions, propel a bird through the air, the cells and molecules making up the bird’s body, too, are propelled, willy-nilly, through the air by the same wants, needs and perceptions. If you add to this the further thesis ... to the effect that these psychological states and processes, though they may ‘emerge’ out of biological and physiochemical processes, are distinct from them, you are apparently committed to the consequence that *these “higher-level” mental events and processes cause lower-level physical laws to be violated*, that the molecules that are part of your body behave, at least sometimes, in ways different from the ways they would if they weren’t part of a living body animated by mental processes” (p. 229).<sup>10</sup>

Steward (2014) says there are two questions that arise from this passage:

- (1) “How [is] the ‘apparent commitment’ to law violation...supposed to arise from the mere assumption that certain psychological states and events (when regarded as distinct from any biological and

---

<sup>10</sup> For more detailed discussion on this quote, see: Steward (2014, pp. 229-232). I employ quite a bit of her argumentative strategy in what follows.

physiochemical ones) can cause the cells and molecules in a bird's body to move?

- (2) "How [is] what follows the italicized phrase...supposed to relate to the italicized phrase itself" (p. 229)?

Recall that law violation is supposed to mean that other, non-physical causes play a role (at least sometimes) in causation at the level of the physical domain. Kim seems to be assuming that biological and chemical factors are clearly in the remit of the physical domain while mental causes are not. If that is right, then Kim's response to (1) would be that that biological and chemical causes can be explained in terms of the physical domain, while mental causes, by definition, cannot. His answer to (2) would be similar. The way what is italicized is related to what comes after is that is that mental states--non-physical entities--play a role in explaining why these biological cells are being propelled through the air.

Steward's (2014) point, however, is that the exponent of downward causation simply does not agree "...that the laws merely of physics are sufficient to dictate in all their myriad details, the movements and changes that will occur as the result of any particular collocation of circumstances" (p. 230). Further, Steward does not think physics is closed even at the level of physics.

Before discussing this, however, I want to point out a few things lest the debate stands at an impasse between the advocate of Causal Closure and an advocate of downward causation. One thing I think can help is that it is not clear that the special domains are explainable in solely physical terms. This is briefly discussed in the next section. Second, the advocate of downward causation usually thinks that the bodily states undergirding mental states play a critical role in their identity; it is just that a mental state cannot be completely identified with the bodily state. This is what would also hold if the domains of, say, chemistry and biology, are not fully reducible to

physics. Clearly, each domain has a lot to do with the physical domain; they are even constituted by fundamental physical entities. It is just that these domains do not fully reduce. The joint insight these two observations offers, then, is that mental states, like biological states and chemical states, are not fully reducible to the physical domain. But that is not to say they do not have anything to do with the physical domain.

At any rate, Steward (2014) says that there is good reason to think that Causal Closure of the physical domain does not hold, even at the level of physics. Steward (2014):

"A free-falling tennis ball, if left to fall by itself, for example, will have to obey Newton's law of gravity. But it can never be predicted, simply on the basis of the law of gravity, that any given individual tennis ball will hit the ground at such and such a velocity given such and such a starting, since the law applies only *ceteris paribus* and hence can be used to predict a precise outcome for a given, individual ball only where no interference is envisaged..." (pp. 230-231).

The thought is that the level of precision often attributed to physical models is due to the fact that such models do not take into account interfering factors, even interfering factors at the level of physics. For example, suppose the tennis ball is on a fixed trajectory. The velocity is set at a certain speed, and the law of gravity is gradually pulling it to the earth. If the speed remains fixed, and the force of gravity remains constant, one could predict with absolute precision where the ball will land. The physical laws in play dictate precisely where the ball will land. But what this model fails to account for is a sudden gust of wind, say, which might blow the ball over a few inches. Or someone might snatch the ball out of the air, mid-trajectory. Or really any number of things could happen to interfere with the trajectory. Once interference is accounted for, it seems that the laws of physics do not dictate certain outcomes.

The Causal Closure theorist has a reply. Steward (2014) puts it like this:

"...it might be retorted that all these various interfering influences must be representable in their turn as physical interventions that can themselves be brought within the purview of other physical laws, so that the *overall* physical outcome will be a result dictated by the entirety of physical laws that apply in any given situation" (p. 231).

The idea is that the tennis ball, by itself, is following a trajectory determined by the laws of physics. The interfering force, say the counter wind, is itself explainable in the laws of physics. And the overall effect, where the ball does land, is fully determined by physical laws active in both events. Any other interfering event can also be so treated. This does seem to be a powerful counter-reply on the face of it. And it is keeping with the historical hope of modern physics, starting from at least Descartes, and peaking in Newton. But Steward (2014) says that in supposing it must be this way, it could be that "we simply let our imaginations run away with us--extrapolating from the closed and ideal systems that physics and mechanics give us to think about to the huge complexity of the real world--in ways that are inadmissible" (p. 231).

Steward does not go into much more detail about the problem, but we have an important clue with the phrase "closed and ideal systems." The problem seems to be that each occurrence is treated as a *discrete event* (*i.e.*, the travelling tennis ball is one discrete event, the gust of wind another discrete event, and so on) in isolation from the surrounding environment, at least in theorizing. By attending carefully to the laws operative in each discrete event, the physicist can determine with precision where the ball will land. That is, unless another event alters the trajectory, such as someone snatching the ball out of the air. But then, the physicist can attend to those forces and, once done the, location of the ball would be precisely known. But again, suppose an explosion displaced the ball after the person caught it (and of course, displaced the person!). And so on. The problem with this view is that it assumes the natural world is

not unlike a well-engineered clock. The movements of such a clock are precisely determined by physical events operating in a law-like fashion (and in this example, in a causation-as-production sort of way). One can attend to the workings of each cog, of the turn of the hands, etc. The physical effects of the watch can be determined by attending to each of the constituent parts and their operation with great precision. That is, the position of the hands, say, can be precisely predicted by accounting for all the relevant events within the watch. However, that this is so is precisely because a watch is a feat of engineering; it is an application of the knowledge of physical laws in such a way that each constituent event operates as close to a *ceteris paribus* situation as possible. The precision of the overall physical picture is the result of the reduction of interference. That is, the different events of the watch are designed to operate without interferences. But why should we think nature operates like a watch, such that each discrete event is free from interference? This is a massive assumption, and one that has come under serious question. It might seem that the natural world is less like a watch, and more like a patchwork of forces in constant flux, interfering with one another. Such a view has been defended by people like Nancy Cartwright (1999):

"The laws that describe this world are a patchwork, not a pyramid. They do not take after simple, elegant and abstract structure of a system of axioms and theorems. Rather they look like--and steadfastly stick to looking like--science as we know it; apportioned at different levels of abstraction; pockets of great precision; large parcels of qualitative maxims resisting precise formulation; erratic overlaps; here and there, once in a while, corners that line up, but mostly ragged edges; and always the cover of law just loosely attached to the jumbled world of material things. For all we know, most of what occurs in nature occurs by hap, subject to no law at all. What happens is more like an outcome of

negotiation between domains than the logical; consequence of a system of order. The dappled world is what, for the most part, comes naturally; regimented behavior results from good engineering" (p. 1).

The hope that everything would be explainable and predictable in physical terms came from the "pockets of great precision." I take it examples of this include the rotation of planets around the sun; the boiling of water at a certain temperature, the measurable and predictable formulations of how fast things fall, and the rate at which they gain speed. But such pockets are quite rare in the sciences. And it is a large theoretical assumption to think the rest of nature will be like these pockets of precision. The tennis ball example includes a variety of seemingly random events that occur, resulting in a lack of precision. And treating each of these various physical events as if they have laws dictating their precise outcomes is, says Steward (2014) "sheer speculation" (p. 231).

That is not to say that the physical laws do not affect physical events. Steward (2014) says: "What the physical laws (at any rate, those that we know of) do...is rather to *constrain* those movements and changes--they place limits on the possible--but they need not be conceived of as dictating every detail of what occurs" (p. 230). For example, given gravity and my own biological limits, I can never jump to the moon unaided. But there is a lot of variety in the distances I can jump; perhaps I manage to jump one inch higher than yesterday due to sheer will and training; or I just have an off day and my jump today was lower than average. Many physical causes and effects are in play, and there is no reason to think that any of them, or combination of them, dictate precisely what the outcome will be. But there are constraints. I will never jump 27 feet into the air unaided, for example.

At any rate, Steward (2014) thinks that there is "...no *scientific* reason to endorse the claim that purely physical laws are sufficient to dictate the movements of every



physical thing" (p. 231).<sup>11</sup> It may turn out otherwise, but until then, we have good reason to deny Causal Closure, even at the level of physics itself. If this is right, then there is also good reason to think, *pace* Kim, that theoretical physics is indeed incomplete.

### 2.1.2 Causal Closure in the Special Sciences

There is a rather large body of literature that argues that certain special scientific domains, such as biology and chemistry, are not reducible to the physical domain.<sup>12</sup> I do not wish here to adjudicate these debates. My interest is in the *possibility* that these domains do not reduce to physics. Of course, it could turn out that these domains are reducible to physics after all. But as of now, this is far from obvious. At any rate, my point is this: if other special scientific domains are not reducible to physics, and causally influence things at the level of physics, Causal Closure fails.

Consider a living cell in, say, a human. Such a cell conducts a great many processes. It conducts cellular respiration, the end result of which is usable chemical energy for the cell. The cell makes proteins for use. Many of the organelles have specialized functions: a lysosome, for instance, breaks down foreign invaders, and it also helps in recycling components the cell is not using. These are only a few examples. Each process is conducted by specialized cellular machinery, so to speak, all of which are ultimately built up out of atoms (usually sharing covalent bonds, but the nucleotides in base pairs in DNA strands, for example, share hydrogen bonds). The behavior of these processes and parts of a cell in a very real way determines the location of certain atoms, as well as how they wind up bonding. For example, the atoms comprising a

---

<sup>11</sup> Steward (2014) also thinks there is no scientific reason to endorse Causal Closure's "indeterministic cousin--that they [purely physical laws] are sufficient, at any rate, to fix the probabilities that any given movement should occur" (p. 231). I take it the reasoning discussed holds *mutatis mutandis*. That is, even if probabilistic causation holds, the various events in nature should not be treated as discrete events in isolation from other discrete events (*e.g.*, the initial trajectory of the tennis ball as one discrete event and the interfering gust of wind as another).

<sup>12</sup> See for example, Henry Stapp (2005) and Robin Hendry (2006).

lysosome are moved from here to there based on cellular commands. A strand of mRNA moves out of a cell and to a ribosome for processing where its instructions result in a protein. The atoms comprising the strand mRNA and the proteins involved in making proteins are all moved in accordance with biological processes and commands. The point is this. If not all of these biological processes are reducible to physics, then biological forces and laws often play an essential role in the physical domain.

If this is right, then once again we have failure of the Causal Closure of the physical domain. In at least some cases what happens at the chemical and biological level determines how lower-level events come to be, such as when a cell determines where certain molecules, and therefore atoms, go. In Kim's terms, there are cases when higher domains inject causal influences into the physical domain.

Nevertheless, there is a residual concern that the 'naturalistically inclined philosopher', to borrow a term from Steward (2014, p. 230), might have. It is the concern that even if biology and chemistry are not reducible to physics, there are still chemical and biological laws that determine what occurs in those domains. The naturalistically inclined philosophers might say that biology and chemistry, say, are still considered 'physical' in some sense, even if not in the sense of 'physics' because of their laws and because of their physical ontology. Biology and chemistry admit of very specific laws and forces that direct the course of events in their respective domains. Biology and chemistry might be viewed as 'natural', perhaps. 'Natural' here can mean having a physical ontology, but operating with other causal influences not recognized in physics (such as those in effect in cellular respiration). But it is difficult to see how mental states could be natural in the same way, such a naturalist might reply, since they are supposed to be able to direct certain physical events without themselves always being physical by definition.

Supposing Causal Closure fails but that biology and chemistry are nonetheless natural, the relevant theses might be re-written to say something like:

Causal Closure\*: If a natural event has a cause at  $t$ , it has a sufficient natural cause at  $t$ ."

Anti-reductionism\*: Mental properties are distinct from, and irreducible to, natural properties.

The Anti-reductionist\* would want to uphold Anti-Reductionism\* for the same reasons she wanted to uphold the old version. If mental states are reducible to some natural state, it would seem that the lower-order natural state is doing the real causal work. Similarly, since natural domains like biology and chemistry have their own laws, these domains' laws can explain their respective events in terms of them<sup>13</sup>. But the Anti-reductionist\* would want to say that at least some mental states are not (fully) the result of a prior natural state. And so, the concern is that, even if we replace Causal Closure with Causal Closure\*, Anti-Reductionism\* still unacceptably entails that there are non-natural causes injecting influence into the natural world.

However, why should the "naturalized" version of Causal Closure be held? What motivation is there? The only motivation that suggests itself is that the other special scientific domains admit of laws and forces that necessitate certain effects. For example, even if chemistry or biology are not reducible to physics, they nevertheless have their own sets of laws. Because they have their own sets of laws, naturalistically inclined philosophers are more ready to accept them as natural domains, even if they have non-physical (physical in the sense of physics) causal powers. The hidden implication is that for something to count as natural, it must have a set of laws such that events can be dictated (entirely) by prior states. But why should we think that laws fix

---

<sup>13</sup> I am discussing laws and law-like forces. But the same point would apply to any probabilistic account of causation.

everything? Indeed, a great deal of work--including that done on probabilistic causation and causal pluralism--has argued persuasively that laws and necessitation are not the only causal shows on the road.

I turn now to causal pluralism and how it can help with the issues at hand. For, even if there are other kinds of causes, and Causal Closure fails, it is important to understand how the causes are related to physical entities, and exactly what role they play in bringing about their effects.

## 2.2 A Causal Pluralist Alternative

The picture now is that it is plausible to think that there are a variety of causes--physical, chemical, biological, mental, etc.--in play that causally influence one another and have roles to play in subsequent events. Moreover, some of these are law-like and influence things in determinable ways, while others do not appear to do so. A good way to start gaining traction on this is to look at the thesis of causal pluralism, which claims there are a variety of different kinds of causes and causal relations. This is what I turn to now.

### 2.2.1 Causal Pluralism in Broad Outline

Jaworski (2011) defines causal pluralism as "the claim that there many different kinds of causes and many different kinds of causal relations" (p. 360). How do we come to discover these? Jaworski (2011) says more:

"To understand these claims [that there are many different kinds of causes and many different kinds of causal relations] it is helpful to appreciate how hylomorphists understand the connection between causation and explanation. According to hylomorphists, causes are explanatory factors.<sup>14</sup> We invariably respond to questions such as *What*

---

<sup>14</sup> Jaworski makes a claim here that not all hylomorphists endorse, or need to endorse.

*caused Zidane's ejection from the 2006 World Cup Final?* or *What caused the Arab-Israeli War of 1973?* by attempting to explain the events in question. That is why explanatory sentences such as (1) can easily be reformulated as causal sentences such as (2) without any apparent loss of meaning:

(1) Socrates died because he drank the hemlock.

(2) Socrates' death was caused by his drinking hemlock" (p. 291).

An explanatory sentence typically features the word "because." What follows "because" is meant to explain why something happened. Jaworski notes that explanatory claims like this can be reformulated as causal sentences without losing meaning. Sentence (2) above makes more explicit the cause of Socrates' death. Jaworski's idea seems to be that for any genuine because statement, a real cause or causal relation is revealed.<sup>15</sup>

Many events have many, many causal factors in play. For example, Jaworski (2011) says different parties might be interested in different aspects of a car crash<sup>16</sup>:

"...A close examination of the [car crash] reveals the following contributing factors among others:

(i) Poor brake design.

(ii) Grading that is insufficient for a sharp curve in the road.

(iii) Inadequate signage warning drivers of the curve.

(iv) The driver's high blood-alcohol level" (p. 291).

How does one go about saying what caused the car crash without unnecessarily proliferating causes in their account? According to Jaworski, different parties--the civil

---

<sup>15</sup> There is of course the issue of when explanations are not "genuine," or do not obtain. The statement--"The duck drowned because Herbert sat down"--does not reveal a genuine cause or causal relation. For more on this, see: van Fraassen (1980b) and Jaworski (2009, 2011). For the purposes of my project, we can make the relatively safe assumption that philosophers and scientists are typically looking to explain what actually happens in the world.

<sup>16</sup> So far as I can tell, Jaworski is paying homage to David Lewis (1986, pp. 215-226) with the car crash example.

engineer, the prosecutor, etc.--are interested in different causes of the car crash for their own purposes. It is not that each one has the one, true cause, or some such. But given their jobs, they have vested interests in particular causal factors and not others.<sup>17</sup> So, while there can be a variety of causes and causal relations in a given context, very often the explanatory interests of the questioner will fix how the question is supposed to be answered.

But it should be emphasized that the causal pluralist thinks that causes and causal relations revealed in real explanations are mind-independent relations between things in the world. They are not just different ways of thinking about those things. The civil engineer, the prosecutor, etc. are all interested in different real mind-independent causal factors.

Causes, for the causal pluralist, can explain effects in many ways and not just causation-as-production ways. For example, a mental cause, if such there be, would not contribute to a subsequent event in the same way as a physical cause; that is, it need not "produce" the triggering of neurons; rather a mental cause might rationalize an action (to borrow Jaworski's (2016 p. 281) use of this term). Here and throughout, let a rationalizing cause be defined as a cause that provides a reason for human action.<sup>18</sup> For example, suppose Morgan is thirsty and decides to get a beer from the fridge. The rationalizing cause for getting up and walking to the fridge is the reasonable desire to quench her thirst. Importantly, I take it that for certain kinds of human action (such as the previous Morgan example), there is a conscious awareness of the reason in the agent. Rationalizing causes will play an important role throughout this project.

---

<sup>17</sup> For this reason, Jaworski (2011) says: "Because causes are explanatory factors,hylomorphists claim that it is often--even typically--unhelpful to speak of *the* cause of something. In any given explanatory context, there are always multiple contributing factors" (p. 291).

<sup>18</sup> This is not to exclude the idea that animal action also employs rationalizing causes. The focus of this project, however, is human philosophy of mind.

Or an angelic cause, if such there be, would not contribute to effects in the same way as a biological cause; perhaps an angelic contributes to what happens by appealing to the conscience in a person, say. Of course, a serious adherent to causal closure would disagree with all this. They would argue "real" causes are causation-as-production style that physics postulates.<sup>19</sup> On this view, there would be no angelic causes or mental causes; unless these factors are actually reducible to physics.

A similar treatment is required for "different kinds of causal relations." According to causal pluralism, a cause can be related to its effect in a variety of ways. For example, a mental cause might have a rationalizing relation to an effect and a physical cause might have a causation-as-production relation to an effect. There could be all sorts of different relations a cause has with an effect.

An example of an opposing viewpoint comes from certain versions of emergentism, according to Jaworski. Jaworski (2016) points out that these kinds of emergentists,

"...assume that there is only one kind of causal relation--the kind that can be described by physics. Even emergent powers are supposed to operate like emergent forces...Consequently, if mental states are to influence behavior, they must do so in the way physical forces do." (p. 280).<sup>20</sup>

The difference between Jaworski's emergentist and causal pluralist is that according to these sorts of emergentists, mental causes behave in a causation-as-production way.

Many who hold this view might also hold that biological causes, if they are emergent

---

<sup>19</sup> Care should be taken here. Causation is a shaky term in physics, and as Steward (2014) has pointed out in a different context, "...fundamental physics, as is often observed, has little use for the concept of causation..." (p. 212). Whether or not forces or force-like entities can and/or should be discussed in terms of causal terms is beside the point here. Whatever causation or forces turn out to be like, the opponent of causal pluralism I have in view is the person who thinks that whatever causation is supposed to be or not be is to be found in the domain of fundamental physics.

<sup>20</sup> Of course, a strict adherent to causal closure would argue similarly--the only causal relation between effects is the kind that can be described by physics.

from chemical causes, behave in causation-as-production ways, and so on.<sup>21</sup> Causes, if they are truly to be causal, must behave in a causation-as-production way. Otherwise, what are they *doing* in bringing about an event?

The causal pluralist can question the assumption that causal *relata* have to *do* (in a causation-as-production way) anything. Events can be contributed to in a variety of ways, and not just in causation-as-production ways. Of course, it is another matter to say how this can be (which is the subject of chapters 3 and 4).

Causal pluralism posits many different kinds of causes and many different kinds of causal relations. I turn now to why the Causal Pluralist must reject Causal Closure.

### 2.2.2 Causal Pluralism and Causal Closure

Merely accepting causal pluralism, however, is only the first step in responding to the exclusion problems because even an exponent of Causal Closure can be a causal pluralist. Here is Kim (2011) again on what Causal Closure does and does not entail, and on what denying Causal Closure would mean:

"Note that physical Causal Closure understood this way does not exclude non-physical causes of physical effects; the closure principle only says that if there are to be such causes, there must also be sufficient physical causes. No non-physical cause of a physical event is essential, and there is no need to go outside the physical domain to give a causal explanation of any physical event" (p. 38).

I take it that when Kim says, "No non-physical cause of a physical event is essential", he means a non-physical cause cannot even be a necessary condition for bringing about a physical event. Causal pluralism and Causal Closure are not on the face of it opposed. Strictly speaking, the exponent of Causal Closure can allow for any number of kinds of

---

<sup>21</sup> Such an emergentist need not be committed to Papineau's Causal Closure adherent who argues that anything made up out of physical entities ultimately operates on other such entities in physical ways. If this sort of emergentist holds that some properties emerge from lower-level bases, they can keep their ontological levels; it's just that each level causes things in causation-as-production sorts of ways.



causes and causal relations. There could be mental, angelic, biological, chemical, and all manner of causes and causal relations. What Causal Closure prohibits is that any of these non-physical causes could be a sufficient or even necessary condition in a physical event.<sup>22</sup> So, again, suppose Morgan is crying, a physical event. Causal Closure requires that there is a sufficient physical cause to explain the crying, and the most likely candidate is her underlying neural state that preceded her crying. There might be a mental cause, or other causes, present in some sense. But these are not sufficient to bring about an effect.

And, as Kim (2011) points out, when Causal Closure is combined with the Causal Exclusion principle, then mental events will always be epiphenomenal with respect to physical events:

"Physical causal closure implies that any putative non-physical cause of a physical event will face competition from a physical event for causal status. The exclusion principle comes into play at this point, yielding the conclusion that in such competition the physical cause always prevails...[M]ental events--and, more generally, all events in the special sciences--are epiphenomenal with respect to physical events" (p. 39).<sup>23</sup>

At any rate, causal pluralism is an unhelpful thesis for our purposes if all it says is that there are a variety of different kinds of causes and different kinds of causal relations. Causal Closure is compatible with such a thesis and as such it is not enough to

---

<sup>22</sup> And according to Kim, there might be a variety of interesting causal relations between non-physical causes and effects. Kim (2011): "Furthermore, as far as closure goes, there may be all sorts of non-physical entities and events, perhaps including mental ones, and there may be all kinds of interesting and significant causal relations among them. The only thing that physical Causal Closure protects is the causal and explanatory self-sufficiency of the physical domain" (p. 39).

<sup>23</sup> Recall that Kim says that the problem with non-physical causes playing essential roles in events is that theoretical physics would be in principle incompletionable since other sorts of causes would be "injecting" causal influence into the physical domain. If other influences are injecting causal influence into the physical domain, then the entities and forces (or force-like powers) postulated by fundamental physics would ultimately be insufficient to explain at least some of what occurs in the physical domain.

provide an alternative picture of mental causation.<sup>24</sup> Consider. Morgan is crying (a physical event). What caused her to cry? Suppose that there was a mental cause--her philosophical angst--in addition to the physical cause of her underlying neural state. Causal Closure will kick in every time and require that if a physical event has a cause, it has a sufficient physical cause. And Causal Exclusion will ensure that her angst is at best epiphenomenal.

To give causal pluralism teeth, we need to adopt a particular version of it; namely, one which holds the further claim that, at least sometimes, non-physical causes are essential or play a direct role somehow in bringing about certain physical effects.

### 2.2.3 Causal pluralism and Causal Exclusion

Denying Causal Closure and holding causal pluralism means that, for at least some physical events, there can be non-physical essential causes in addition to physical causes. However, it would be a hollow victory indeed if causal pluralism required systematic overdetermination across the board. This section addresses Causal Exclusion and why causal pluralism is compatible with it.

Causal Exclusion says that no event can have more than one sufficient cause unless it is a case of genuine overdetermination. The causal pluralist is free to retain this thesis, and she probably should. Otherwise, systematic overdetermination appears in many unwelcome places. One of the worst places for systematic overdetermination to show up is in philosophy of mind. Consider intentional actions. For example, Morgan decides to throw a baseball, locks back her right arm, and throws a wicked fastball.

Jaworski (2016) explains:

---

<sup>24</sup> And recall from chapter 1 that rejecting the Causal Exclusion principle does not get us very far either. If Causal Closure still holds, then caused physical events still have a sufficient physical cause, and it would appear that other causes roles in bringing things about is more mysterious than ever. But, Causal Closure is a good principle: we do not want to multiply sufficient causes needlessly, whether or not one is a causal pluralist.

“Rejecting [Causal Exclusion] ...implies that [intentional] actions are overdetermined—that each has both a sufficient physical cause and an independent sufficient mental cause, and this too has awkward implications. In particular, it implies that the action could have occurred without a mental cause or without a physical cause. If certain physiological events were sufficient by themselves to bring about a particular action, then no mental cause was necessary for that action to occur: the action could have occurred without any mental cause at all. If, on the other hand, the mental cause by itself was sufficient to bring about the action, then no physical cause was necessary for the action to occur. The action could have occurred without any physical cause at all. Yet, it seems implausible to suppose that an intentional action could occur without both mental and physical causal factors. It seems implausible to suppose that I could throw a baseball without certain events in my nervous system bringing about my limb movements, and it seems equally implausible to suppose that those limb movements could count as an intentional action without the relevant intentions, beliefs, or desires” (p. 276).

Denying Causal Exclusion would mean that such events are overdetermined. In the case of Morgan, the action of throwing a (wicked) fastball is overdetermined. This means that if the decision to throw had obtained, the fastball would have been thrown *without* physiological activity in the brain, nervous system, or muscles! That is, the decision would have sufficed without a physical base. If the baseball throwing is overdetermined, it means it would have occurred if a mental cause was present, but no physical cause; and it would have occurred if a physical cause was present, but no mental cause.

Moreover, while it is a strange and unwelcome result that the baseball throwing could occur without physical cause(s), it is also strange and unwelcome to suppose that this sequence of physiological causes could count as a coordinated intentional action without a mental cause. Throwing a baseball correctly does not seem to be the result of unintentional bodily movements on a par with spastic arm movement, say. Throwing a kind of baseball pitch requires skill: getting the ball to go the right speed and to hit a small target (the catcher's glove). An application of skill seems to require some kind of mental state, such as intention.

Causes, therefore, are best thought of as 'contributing factors.' Mental and physical causes contribute to the ball throwing in different ways. What are these different ways? This is the subject of the next subsection.

#### 2.2.4 Rationalising vs. Triggering Causes

Using Jaworski's (2016) terminology,

"...beliefs and desires cause or contribute to actions in one kind of way-- they *rationalize* actions--and neural events cause or contribute to actions in a different kind of way--they *trigger* the muscular subsystems involved in actions" (p. 281).

According to Jaworski, contributing factors are not competing for status of sufficient cause of an event since each one contributes to what happens in different ways. For example, in the case of philosophy of mind, a rationalizing cause and a triggering cause are not sufficient by themselves to bring about an action. To compete for status of sufficient cause would require that either the physical cause, or the mental cause, to bring about the effect *on their own*.<sup>25</sup>

---

<sup>25</sup> Later, we will see that Jaworski's neutrality on Causal Closure contradicts the idea that rationalizing causes can contribute to physical action.

Nevertheless, the mental cause and physical cause are each essential to the subsequent effect. This can be expressed in conditional statements. Had the rationalizing cause not occurred, Morgan would not have thrown the baseball. Had the triggering cause not occurred, Morgan would not have thrown the baseball. Neither of these statements require the implication that had either the rationalizing cause or the triggering cause happened on their own, Morgan would have thrown the baseball. One plausible way of looking at the rationalizing and triggering causes is as necessary conditions of the subsequent event; neither are sufficient on their own.<sup>26</sup>

### 2.3 Lingering Worries

Although denying Causal Closure and embracing causal pluralism are critical in solving the exclusion problem, it is not enough. While Jaworski's rationalizing and triggering cause move is compelling, there are two lingering worries. For one thing, even if Jaworski's initial account explains how a rationalizing cause and a triggering cause are related to their effects, it is unclear how they are related to one another. However the rationalizing cause is related to an effect, it is also related somehow to the underlying neural states that also gave rise to the same effects. What is the nature of this relation? Secondly, for all Jaworski has said, it is not clear why the triggering cause is not sufficient for subsequent physical effects. The neural events involved in Morgan's baseball throwing still have the effect that Morgan raises her hand, Morgan throws the ball, and so on. We need an account of why the triggering cause is not sufficient for the effects in question.

---

<sup>26</sup> It is a further question of whether or not they are jointly sufficient. The most obvious reading would be yes, they are jointly sufficient. However, this closes the door to the possibility of their being other kinds of cause that might be in play. A good middle ground is to say that citing the rationalizing and triggering causes are jointly sufficient for the purposes of most kinds of explanations of actions. "Explanations" here should be understood as referring to the explanatory interests of someone asking a question like: "What caused Morgan to throw the ball?" Citing the triggering and rationalizing causes will likely satisfy the person's interests (and in many cases, citing the rationalizing cause alone might satisfy a questioner). But such an answer still leaves open the possibility of other kinds of causes if such there be.

In what follows, I am going to look at each of these questions in greater detail. Following this, I offer a diagnosis of why Jaworski might not have addressed these issues. In short, it is because he wants to remain neutral on the Causal Closure principle.

### 2.3.1 How Are the Rationalizing and Triggering Causes Related to Each Other?

It is unclear how the rationalizing and triggering causes are related to each other on Jaworski's account.<sup>27</sup> As will be seen, the problem is that the structure of neural firings is *included in the triggering cause* on his account and this makes it mysterious how the triggering cause could be related to the rationalizing cause.

Jaworski clearly wants to hold some version of Anti-reductionism. A rationalizing cause rationalizes an action, while a triggering cause sets the muscular subsystems in motion. Recall that Jaworski (2016) described rationalizing and triggering causes this way:

"Beliefs and desires cause or contribute to actions in one kind of way-- they *rationalize* actions--and neural events cause or contribute to actions in a different way--they *trigger* the muscular subsystems involved in actions" (p. 281).

Let's shift the example from the baseball throwing to another kind of example. Morgan, tired from philosophy of mind class, decides to take a Saturday off. She decides to play a video game she has never played before. The game turns out to be extremely engaging and she is really enjoying herself. However, eventually, certain biological conditions cause her to be thirsty. She desires a drink. But Morgan is also really enjoying her game. In her thirsty state of desire, she weighs the pros and cons of getting up right

---

<sup>27</sup> Presumably, one could set up a how-question:

(C) How are rationalizing causes and triggering causes related?

One might then supply an answer which reveals the causal relation based on the logic of how-questions. But however that works out, the answer will imply a certain metaphysic about how these sorts of causes can be related. Even if Jaworski's favored version of causal and explanatory pluralism reveals a variety of causes and causal relations, some contexts require a worked-out metaphysic in order to explain anything. For an overview of the logic of answering how-questions, see: Jaworski (2011, pp. 292-296). For an analysis of how- and why-questions, see: Jaworski (2012).

now. Eventually, she decides to stand up and fetch for herself a Pabst Blue Ribbon beer from the fridge. When we ask: Why did Morgan get up from the couch? We could give two "because" formulations:

(A) Because Morgan was thirsty and wanted a drink.

(B) Because certain neural events triggered certain muscular subsystems.

Each can be reformulated as causal sentences:

(A\*) Morgan's desire for a drink caused her to get up.

(B\*) Certain neural events caused certain muscular subsystems to be triggered.

While this set up seems to help with an important aspect of mental causation—namely, how causes are related to effects—it does not help with understanding how exactly the triggering and rationalizing causes are related to one another.

The problem is that the triggering cause includes the coordinated sequence of events that makes up Morgan's walk to the fridge (and indeed the sequence of events that leads to the pitcher throwing the ball in such-and-so a way). It would appear, then, that the walking to the fridge (or the throwing of the ball) could occur without any kind of rationalizing cause at all. This clearly has implications for how the rationalizing cause is related to the subsequent events of going to the fridge. This is the subject of the next section. But there is also a problem as to how the rationalizing cause is related to the structure of these events themselves. Jaworski has an account of how this works, but, as shall be seen, including the structure of events in the triggering cause still creates problems.

Jaworski can easily augment this sketch with the hylomorphic picture he paints throughout his book, especially in Jaworski (2016, pp. 155-162). Jaworski describes what he calls "activity making structures" and "activity composition." Jaworski (2016) defines activity making structures:

"The idea there are activity-making structures is based on the observation that the activities of structured individuals involve coordinated manifestations of the powers of their parts. Consider the activities we humans engage in: walking, crawling, driving...Each of these activities involves some of our parts manifesting their powers in an organized or coordinated way. When we perform these activities, we impose an organization or order on the ways those parts manifest their powers...[S]ome structures are dynamic...They are not relatively stable spatial relations among an individual's parts. They instead have a temporal dimension: they comprise sequences of changes over time" (pp. 155-156).

Many such activities are voluntary, such as the ones in the list: walking, crawling, driving. And we can throw in Morgan's getting up from the couch as being a dynamic structured event. Other activities are involuntary and still dynamically structured and so are instances of activity-making structures: Jaworski (2016) cites the Krebs cycle as an example, wherein a series of processes occurs that results in usable energy for the cell (pp. 155-156). We could also include other intercellular activities such as glycolysis (and photosynthesis in plant cells). The thought is the same: activity-making structures are responsible for imposing organization and order through a manifestation of certain powers of certain parts, whether or not these structures are engaged in deliberately (crawling, throwing baseballs, going to the fridge). Jaworski (2016) gives this definition of activity composition:

"Here is one attempt to define a notion of activity composition: Suppose that  $a$  is a structured individual with the power to engage in activity  $A$ . Suppose, moreover, that  $b_1, b_2, \dots, b_n$  are individuals, and that a subset of the  $b$ s are proper parts of  $a$ . Suppose, finally, that each  $b_i$  has the power



to engage in an activity,  $A_i$ :  $b_1$  has the power to  $A_1$ ,  $b_2$  has the power to  $A_2$ , and so on. We can then define a notion of activity composition as follows:

*Activity composition*: Necessarily,  $a$ 's  $A$ -ing at  $t$  is composed of  $b_1$ 's  $A_1$ -ing at  $t_1$ ,  $b_2$ 's  $A_2$ -ing at  $t_2$ ,...and  $b_n$ 's  $A_n$ -ing at  $t_n$  exactly if  $b_1$ 's  $A_1$ -ing at  $t_1$ ,  $b_2$ 's  $A_2$ -ing at  $t_2$ ...and  $b_n$ 's  $A_n$ -ing at  $t_n$  are structured  $A$ -wise, where time  $t$  includes the times  $t_1, t_2, \dots, t_n$  (p. 160).

Morgan, a structured individual, employs an activity-making structure when getting up from the couch. Getting up, an activity Morgan has the power to engage in, requires a coordinated sequence of events which she sets in motion via her reason, and also the triggering, mechanistic causes that power the getting up, as it were. Morgan's getting up is composed of her relevant parts getting up at time  $t$  exactly if the relevant parts are getting up by being structured in such a way as to constitute an instance of getting up.

The question is: is Morgan getting up more like the Krebs cycle or does it include deliberation? And if it is more like the latter, what role does deliberation play in the event? Yes, the answer is supposed to be that it rationalizes the action (provides a reason for getting up), but for all that has been said it is unclear what deliberation amounts to.

Is rationalizing one kind of dynamic structure that includes the certain coordinated manifestation of parts? If so, which parts? It does not seem like the parts that are involved in actually moving can be included because Jaworski identifies the neurons and muscular subsystems as part of the triggering cause. Does rationalizing involve certain parts that eventually *set in motion* the triggering causes? If so, this could explain how a rationalizing cause is related to a triggering cause; it sets it in motion. But now it is mysterious: which parts of Morgan are involved in rationalizing? And

wouldn't these parts include her biology in some way, like say certain neural patterns that underlie her rationalizing? Then it seems like the patterned sequence of neural firings are the triggering cause in accordance with Jaworski's definition of the triggering cause.

However, this does not seem right; recall that Jaworski wants the triggering and rationalizing causes to run in parallel, each explaining an aspect of subsequent action; rationalizing is going on while triggering is going on. It does not appear to set it in motion. So what way is the rationalizing cause related to the simultaneous triggering cause?

A problem with sufficiency looms too. If the triggering causes explain the sequenced movements of Morgan and the baseball pitcher, what causal work is there for the rationalizing cause? This is the subject of the next sub-section.

### 2.3.2 Why Don't These Two Kinds of Causes Compete for Sufficiency?

Recall that Jaworski explained the relevant kind of sufficient causation through the example of a baseball player throwing a pitch. For a mental cause to be sufficient in the right way, it would have to be able to bring about the pitch even without a synchronous physical cause (neural event triggering muscular subsystems) occurring. Jaworski (2016):

"It seems implausible to suppose that I could throw a baseball without certain events in my nervous system bringing about my limb movements, and it seems *equally implausible* to suppose that those limb movements could count as an intentional action without the relevant intentions, beliefs, or desires" (p. 276) (italics mine).<sup>28</sup>

---

<sup>28</sup> I am sympathetic to Jaworski's view about how implausible it seems the baseball throw could happen without an intentional state, but it is not obvious that this is as implausible as the mental cause bringing about the throw without certain events in the nervous system. In fact, because the incongruity between these two is so striking, I think we are owed more of an explanation as to why they both seem equally implausible. As I argue later, Aristotle's metaphysics of causation helps with this.

Even though it does seem implausible that the baseball could be thrown without an intentional state, it still seems like the triggering causes are in direct competition with the intentional (rationalizing, in this case) cause. Here is why.

Suppose that the pitcher's nervous system was wired up to a certain kind of technology. This technology could be controlled by remote control by the scientist who invented it. It gives a series of inputs into the nervous system that result in action, say, the pitcher throwing a baseball. The pitcher stands there, with no real intentions about anything formed. Suddenly, the scientist fiddles with the remote control and the pitcher throws the ball, still in an intention-less state. It seems like the triggering causes involved in getting the muscular subsystems going are *sufficient* to explain the throw. Why is this not also the case in the normal scenario where the pitcher throws the baseball, intention and all?

It seems like the same triggering causes are in play in the normal case and so are in direct competition with the rationalizing cause (the pitcher's intention). They are in competition because the triggering causes appear to be sufficient for the action, including even the structured coordination of the various systems and subsystems. What exactly does the intention contribute to the subsequent action not already "covered" by the triggering causes?<sup>29</sup>

What we need is a more detailed account of how exactly the rationalizing cause is related to the effect in such a way that it is not pushed out by the triggering cause. I certainly think this is possible to do. One somewhat mysterious clue that this is the case comes from the neurosurgeon Wilder Penfield's (1975) work. Penfield (1975) set up an experiment with patients very similar to the baseball pitcher:

---

<sup>29</sup> It is for this reason that it is not clear to me why Jaworski's thinks that, when considering intentional action, it is equally implausible that such an action could happen without the nervous system and the intention itself.

"When I have caused a conscious patient to move his hand by applying an electrode to the motor cortex of one hemisphere, I have often asked him about it. *Invariably*, his response was: 'I didn't do that. You did.' When I caused him to vocalize, he said: 'I didn't make that sound. You pulled it out of me'" (p. 76) (italics mine).

In both cases--moving the hand and vocalizing--Penfield brought about the triggering causes that resulted in the hand moving or the instance of vocalizing. And what is of great interest is that *every time*, according to Penfield, the patient knew the action did not originate with his own intention but was the result of Penfield's brain meddling. Jaworski (2016) uses this same example from Penfield to back up his point that it seems equally implausible to think intentional actions could occur without intention (p. 276). A very plausible reason why the one action is different than the others is the intentional state. But much more is needed to spell out the relation between the intentional action and the subsequent event.

Indeed, on the face of it, Penfield offers good evidence that intentionality has some kind of rationalizing contribution to an effect: the agent was aware that the movements from brain stimulation did not originate from himself.<sup>30</sup> But it is not enough to merely point at Penfield's example; what is required is a theory that shows how intention (or rationalizing of any sort) is related to an effect. This is what I think Aristotle offers, but before I turn my attention to chapter 3, I would like to discuss why I think Jaworski doesn't address these points in more detail. The key issue, to my mind, is that he seems to try to remain neutral on Causal Closure of the physical domain.

---

<sup>30</sup> It might strike the reader as interesting that Penfield held a certain version of substance dualism (see: Jaworski (2011, p. 376).

## 2.3.3 Jaworski's Neutrality on Causal Closure

Why has Jaworski not provided more detail about how the triggering and rationalizing causes are related to each other, or how the rationalizing cause is related to an effect? The key reason, I think, is that Jaworski wants to leave it an open empirical question whether or not causation is closed under physics.

That Jaworski is committed to a neutrality on Causal Closure can be supported in the following way. Jaworski (2011) says: "The existence of other causes other than those described by physics...does not imply that physical laws must be violated..." (p. 120). It is not clear what Jaworski means by "violated." In context, he is directly addressing David Lewis' argument for the identity theory of mind, which employs a version of the Causal Closure thesis. Jaworski (2011) interprets Lewis' version of Causal Closure thus: "...physics is capable in principle of providing an exhaustive description and explanation of all causes" (p. 117). This is further supported by Jaworski's (2016) comment elsewhere: "...the behavior of structured individuals never violates the laws governing their fundamental physical constituents" (p. 161).<sup>31</sup> It seems, then, Jaworski has in view that a violation of physical laws would be in an instance where a non-physical cause influences the physical domain in some way. And so, he appears to interpret Causal Closure in a very similar way to Kim, where a violation of physical law is "injecting causal influence" into the physical domain.

It is this neutrality (and possible embrace of Causal Closure) that I think is the reason Jaworski does not provide more detail about the relevant relations. Recall the baseball pitch. At time  $t_1$ , the pitcher's hand is in her mitt; at  $t_2$  her hand is behind her head in a particular form; at  $t_3$  her arm is bent in front of her and the ball is released. Suppose we consider the location of the fundamental physical entities at each of these

---

<sup>31</sup> Much of this discussion here is influenced by insights from William Simpson in addition to the material from Steward (2014) and Cartwright (1999, 2002) discussed earlier in this chapter.

times. We might ask: why are these fundamental physical entities at this location at  $t_1$ , then  $t_2$ , and finally  $t_3$ ?

If part of the reason the fundamental physical entities are located at different places involves the pitcher's intention to throw a slider, then it looks like we have a violation of Causal Closure since a psychological reason would be influencing the physical domain.<sup>32</sup> On the other hand, suppose it turns out to be the case that physics is not causally closed. Then there would be no problem here. It just turns out that other domains, such as the psychological, causally influence the physical domain. I think Jaworski's attempt at neutrality on this issue is well-intentioned. But a decision is required on Causal Closure: either non-physical mental states make a non-overdetermined impact on the world, or they do not. And if they do impact the physical domain, how (the subject of chapters 3 and especially 4)?

Suppose Causal Closure holds. This has serious implications for both the relation between rationalizing and triggering causes (the subject of the previous section) as well as their effects. As to the latter, it seems to rule out a rationalizing cause having any causal power to bring about an effect; the triggering causes move the body in the coordinated fashion required. This is so because any physical effect, if it has a cause, has a sufficient physical cause. And we are back to the exclusion problem discussed in chapter 1. And if this is the case, saying that mental states "rationalize" the action is of no help on its own. What would this mean? What difference would that make?

We saw earlier that Jaworski, as a hylomorphist and causal pluralist, is committed to structures being ontologically basic, and to many different kinds of causes and causal relations. But even if one is an advocate of these views, Causal Closure of

---

<sup>32</sup> There might be an alternative explanation wherein the intention itself is composed of physical properties, such as neural patterns, which in turn are made out of atoms, etc. But this violates the principles of Jaworski's (2016) hylomorphism because activities like baseball throwing are "structured activities" and structures are ontologically basic (pp. 155-177).

the physical domain causes problems spelling out key causal relations. For Causal Closure does not rule out other causes, or even structure; it just commits one to the thesis that physical events, if caused, have a sufficient physical cause.

If hylomorphism and causal pluralism are to be of any help, we need to scrap Causal Closure of the physical domain. I take it that a key motivation behind hylomorphism, whether Aristotle's, Jaworski's, or anyone's, is to explain how higher-order domains can have a causal impact on the physical domain. And these causal impacts are partially explained by the ways certain physical quantities are structured. Take, for example, a mousetrap. None of the five constituent parts have the capacity to catch mice on their own, but only after a certain configuration is achieved. A mouse tries to get the cheese and the hammer cracks its neck. The reason the fundamental physical particles of the hammer are now in *this* location (*i.e.*, cracking the mouse's neck) and not in *that* location (*i.e.*, in a position ready to break the mouse's neck) are partly due to biological structures (the mouse in action and located at the striking point), and the higher-order structure of the mousetrap (the configuration of the five pieces).

But denying Causal Closure only opens the door for explaining how psychological states can be causally efficacious. Work is required to show how this can be so.

#### 2.4 Summing Up and Some Questions for Jaworski

At this juncture it is time to summarize. Although historically, Causal Closure of the physical domain has enjoyed a privileged status, there are now very good reasons to think that it is undermotivated. For one thing, there are a variety of debates on whether or not the scientific domains are reducible to physics. For another thing, as Steward points out, it is not clear that physics is causally closed, even at the level of physics, as her tennis ball example demonstrates. And Cartwright makes a strong case for the dappled world. None of this is definitive argumentation against the Causal Closure

theorist, but there is enough argument to call it into question and to theorize as if it were not true.

I also suggested that causal pluralism is a helpful way forward and I set up Jaworski's version of causal pluralism as a first pass. A key issue with it is that, while it is clearer how certain causes are related to effects, it's often less clear how causes are related to each other. This is especially problematic in philosophy of mind, where the relation between rationalizing and triggering causes is pressing.

While the rationalizing and triggering causes provide a good model for a way forward, Jaworski's neutrality on Causal Closure creates problems. It becomes a pressing matter to answer these questions:

- (1) Why don't the rationalizing and triggering causes compete?
- (2) How are the rationalizing and triggering causes related?

In what follows, I am offering Aristotle's causal pluralism as an alternative. As we shall see, Aristotle's causal pluralism has a built-in metaphysics that, when applied to Jaworski's rationalizing and triggering causes (the subject of chapter 4), shows how they are related. In the next chapter, I explain Aristotle's four causes, why the causes do not take over the explanatory work of one another, why they require one another, and how each of the causes is related to one another.



## Chapter 3: Aristotle's causal pluralism

The previous chapter ended with some difficult questions about why the rationalizing and triggering causes do not obviate the need for each other. Further, if they do not obviate the need for each other, how are they related?

Mapping rationalizing and triggering causes onto Aristotle's four causes answers these difficult questions. The key aim of this chapter is to explain how Aristotle's causal pluralism works, why the four causes do not compete with each other, and how they are related in cases of non-mental causation. The final chapter will take up the task of providing these explanations in cases of mental causation, as well as the task of mapping rationalizing and triggering causes onto Aristotle's framework, and explaining how this helps us answer the questions we were left with at the end of chapter 2.

This chapter proceeds in the following way. First, I discuss Aristotle's account of the relationship between causation and explanation. Next, I introduce the four causes--Material, Formal, Efficient, and Final. After this, I discuss why the causes do not compete with one another. That is, I discuss why each has a particular job to do that cannot be done by the other causes. A key clue to this is that each cause answers different questions, and, upon analysis, each of the causes seems to require the presence of the others (in certain cases). Finally, I discuss the metaphysical relations between the causes, focusing in particular on the difficult question of the relationship between material and formal causes.

### 3.1 The Relation between Explanation and Causation on Aristotle's View

A causal pluralist should be distinguished from the explanatory pluralist. A causal pluralist tends to think that different kinds of causes and causal relations obtain are features of the mind-independent world as opposed to just ways we think about the world. An explanatory pragmatist tends to think that the way different causes feature in explanations are just ways we tend to think of the world, or according to Bas van

Fraassen (1980b), ways in which we answer how and why questions (p. 127). Whether or not these causes and causal relations are mind-independent features of the world is another matter. The key difference between these two approaches is that the causal pluralist thinks that the variety of causes and causal factors revealed in (good) explanations reveal mind-independent causes and causal relations in the world, whereas the explanatory pluralist is at least agnostic on the point that (good) explanations reveal such mind-independent features.<sup>33</sup>

There are good reasons to think that Aristotle is a causal pluralist and that the four causes he discusses are meant to be taken as mind-independent features of the world. In what follows, I am going to briefly discuss the history of Aristotelian exegesis on this point and then explain why I think it is very plausible he took himself to be discussing mind-independent causal features. Again, my aim is to provide a plausible view of Aristotle's work in service of solving a contemporary puzzle; I am not here going to adjudicate between long-standing exegetical debates.

In the 90s, pretty much everybody thought that Aristotle was an explanatory pragmatist. Cynthia Freeland (1995) says: "Among Aristotle scholars current opinion is virtually unanimous that theory of the four causes is really a theory of explanation" (p. 49). Indeed, some like Bas van Fraassen (1980a, 1980b) believed Aristotle was the forerunner of explanatory pragmatism. Freeland goes on to present a convincing series of arguments as to why we should take Aristotle to be a causal pluralist, which I warmly commend to the interested reader.<sup>34</sup>

---

<sup>33</sup> A paradigmatic example of "good" explanations are scientific explanations wherein a particular feature of the world (say, how genes go from coded base-pair sequences to the formation of a protein) is explained through observation and experiment. The explanations provided are meant to give an account of what actually happens. At any rate, whether explanatory pragmatist or causal pluralist, everybody has the same concerns about good vs. poor explanations.

<sup>34</sup> One way Freeland (1995) argues for this is by discussing Aristotle's views on explanation in the *Posterior Analytics* and the *Physics*, concluding that the pragmatist interpretation gets things wrong. However, for my purposes, it will be enough that there is a growing, robust literature on Aristotle's causal pluralism.

Freeland also diagnoses why so many chose to interpret Aristotle as an explanatory pragmatist. A big part of the reason was that many thought that at least three of the causes could never be "causal" in the way we are used to using the term.

Freeland (1995) says

"The motivation for the explanation approach seems mainly negative. Advocates of this approach deny that Aristotelian αἰτίαι are causes because they take it as obvious that (at least) three of the Aristotelian αἰτίαι differ quite radically from causes as 'we' conceive them (necessary or sufficient conditions, events constantly conjoined to effects, events probabilistically conjoined to effects, or others). On the explanation approach we can preserve some parallels between Aristotle's and 'our' accounts of causes, provided we reserve use of the term 'cause' for what he calls moving or efficient causes. Then presumably the other αἰτίαι (formal, final, and material) fall into place in this picture by being interpreted as explanatory factors relative to this one 'real' (i.e. efficient) sort of causal relation that seems most to resemble our own..."

(p. 50).

I would like to discuss this diagnosis in light of the talk in chapter 1. There, I argued that a lot of the historical baggage surrounding Causal Closure had to do with causes often being thought of in the causation-as-production way. I think this had a lot to do with the way Aristotle was interpreted in the 90s (and indeed several decades before). The only kind of cause in his four-causal program that looked anything like causation-as-production was the efficient cause.

This sort of reasoning seems very close to the surface, for example, in Terence Irwin (1988):

"If a cause of one type (e.g. a substance) is reducible to a cause of another type (e.g. an event), it is worth considering whether some of the four causes may not be reducible to others. We may say that the matter and the form are causes of the statue; but 'the statue' seems to be an inexact description of what is to be explained. A more exact description asks for the explanation of the statue's being made, or of its being as it is, or of its remaining in existence. Once we have specified the effect more fully, we also want a fuller specification of the cause. *Simply citing the form or the matter does not seem to provide the proper explanation; 'the form is the cause of the statue' seems a mere abbreviation. The cause of the fully specified effect is properly and fully described as 'the sculptor exercising her craft' or 'the customer giving her order'. But now we have stated the efficient cause and not simply the form or the matter"* (pp. 95-96) (italics mine).

In this chapter and chapter 4, I discuss in much more detail the causal picture given by taking Aristotle as a causal pluralist. But here, we can see that Irwin seems to be working with a causation-as-production view. He says that matter and form "cause" the statue, but this is inexact. The more exact specification of cause is what brought about the statue. However, this is only imprecise if Aristotle wants form and matter to be understood in a kind of causation-as-production sort of way. The upshot is that because final, formal, and material causes are in no way "productive," they cannot be truly causal; at best they are short-hand for the efficient cause.

As I shall argue later, this is not the right way to understand what form and matter are doing in a causal explanation. Matter and form explain what things are changed from and into when an efficient cause is acting (for example, a change from a pile of wood, nails, and plaster and into a house).

Causal pluralists are serious about there being a variety of different sorts of causes and causal relations, not just causation-as-production. Of course, the question is: was Aristotle a causal pluralist of some kind? If so, what are the reasons for thinking so?

I won't dwell too long here. I will argue that the most straight-forward way of understanding Aristotle's introduction of the four causes is to take them at face value. After this, I will compare my reasoning to Christopher Shields (2014) who discusses the same passage. It is a very plausible reading of Aristotle that he was a causal pluralist.

Aristotle introduces the discussion of the four causes this way:

"Since the object of our enquiry is knowledge, and we do not think we know a thing until we have grasped why [διὰ τί] it is so (where this is to grasp its primary cause [πρώτην αἰτίαν]), it is clear that we must also find this in the case of coming to be, perishing, and of all natural change, so that when we know the principles of things we can endeavour to refer what we are seeking back to these principles" (*Phys.* 194b16-23)  
(translation Shields (2014, p. 48)).

It would appear that Aristotle is genuinely interested in explaining the causal principles of natural change. And once these principles are grasped, we can explain any changes in nature through them. The most straightforward reading of this text is that the following principles Aristotle gives are to be taken as mind-independent causal factors.<sup>35</sup>

Christopher Shields (2014) argues along similar lines using the same passage.

Shields (2014) says that:

---

<sup>35</sup> Of course, not everyone agrees. Many who thought he was an explanatory pragmatist understood this passage differently, often comparing and contrasting it with his different account of causes in the *Posterior Analytics* (for discussion on this, see Freeland (1995)). Others thought that because he stayed so close to common beliefs about causation, he couldn't be seen to engaging the subject with the objective rigour it needed. Take for example Irwin (1988): "His actual doctrine of the four causes is too purely dialectical, and stays too close to common beliefs, to make clear the centrality of the efficient cause; and his failure to make this clear seriously affects some of the argument in Book ii" (p. 96). At any rate, I do not mean to suggest that my reading of the passage is the only reading.

"An explanation requires grasping the primary cause of what we want explained; something's primary cause is not, however, something whose mention merely happens to satisfy someone seeking an explanation. Rather, a primary cause is what in fact makes it the case that a certain state of affairs obtains...It is important, then, that we reflect upon what makes a connection between states of affairs suitable to ground an objective explanation. Aristotle contends that genuine connections, the sort cited in objective explanations, are causal" (p. 49).

Shields thinks that Aristotle's preferred kind of explanation is objective as opposed to subjective. An objective explanation, says Shields (2014), occurs when:

"...*x* explains *y* just in case (i) *x* and *y* are states of affairs in the world, and (ii) states of affairs of the *x*-type *cause* states of affairs of the *y*-type" (p. 48).

An objective explanation gives the causal account from *x* to *y*.

At any rate, Shields is not the only one who considers Aristotle to be a causal pluralist.<sup>36</sup>

In this section, I have looked at how Aristotle was interpreted as an explanatory pragmatist for a variety of reasons. However, there are good reasons to think that he is a causal pluralist and thinks the four causes are actually mind-independent causal principles that are features of the world. I did not adjudicate between any of these debates because all I need is a *plausible* interpretation of Aristotle. In what follows, I explain how the causes work and how they are related to one another and their effects.

---

<sup>36</sup> For a few examples, see: Christopher Shields (2014), Cynthia Freeland (1995), and Andrea Falcon (2015).

### 3.2 Aristotle's Four Causes

In this section, I introduce each of Aristotle's four causes. He introduces them here:

"With these distinctions in place, we should look into the question of how many causes [τῶν αἰτίων] there are, and what they are like. For the point of our investigation is to acquire knowledge, and a prerequisite for knowing anything is understanding why [διὰ τί] it is as it is—in other words, grasping its primary cause [λαβεῖν τὴν πρώτην αἰτίαν].

Obviously, then, this is what we have to do in the case of coming to be and ceasing to be, and natural change [τῆς φυσικῆς μεταβολῆς] in general. Then, once we know the principles [τὰς ἀρχὰς] of these things [of coming to be, ceasing to be, and natural change], we can try to analyse anything we are looking into in terms of these principles”

*Phys.* 2.3 194b16-23 (translation Waterfield's (2008)).

Study of the causes--how many there are and what they are like--will furnish us with the causal principles to analyze the world of natural change. Natural change [τῆς φυσικῆς μεταβολῆς] comes in all sorts of ways: why things--like fish and statues--come to be and pass away; why rainstorms occur; why trees grow leaves, etc. What Aristotle is aiming at is no less than the causal principles operative in the world such that any natural change whatsoever can be explained in terms of them.<sup>37</sup> He also applies these

---

<sup>37</sup> I am clearly interpreting Aristotle as a "realist" about the causes. That is, Aristotle is not focusing on the pragmatics of explanation, *pace* van Fraassen et al. Rather, his aim is to elucidate real principles at work in nature. Again, the aim here is to present a plausible interpretation of Aristotle's thought, not to settle any exegetical debates. For literature on those who consider Aristotle an explanatory pragmatist about the four causes see: Julius Moravesik (1974), Martha Nussbaum (1978), and Julia Annas (1982). For a more detailed list, see Freeland (1995). For those who interpret Aristotle as a realist about the causes, see: Christopher Shields (2014), Cynthia Freeland (1995), and Andrea Falcon (2015).

principles to mental causation, as we shall see in chapter 4. At any rate, he identifies four causal principles, which I discuss next. First, a note on how Aristotle uses the word "cause."

Aristotle is using a wider notion of "cause" [ἡ αἰτία] than we are accustomed to. For example, Aristotle considers the material something is made up out of as a kind of cause, such as the bronze of a statue. This strikes many as implausible. Why? Because we tend to think of causation as having "oomph," to borrow a turn of phrase from Thomas Kroedel (2015). "Causal oomph" is a way of expressing the "causation-as-production" view of causation. Recall that causation-as-production "... involves continuous spatio-temporal processes and the transfer of energy between cause and effect" (p. 369). Causation is thought to involve forces or force-like powers operating on something. I am not saying that everyone thinks of causation in these terms, but it has enjoyed a place of prominence for a long time and is frequently treated as the "go-to" for defining or specifying the term 'cause'.

'Cause' has a wider scope in Aristotle's thought. The word ἡ αἰτία means cause, but also includes connotations of "being responsible for". The term in context could be understood as finding those principles "responsible for" why something is as it is. The bronze is "responsible for" the statue in that it is the material that became a statue. This particular statue would not be here without it.

With these points in place, I now turn to introducing the four causes in turn. I borrow an extended example of coming upon a bronze statue deep in the mountains from Shields (2014, pp. 50-52).

### 3.2.1 The Material Cause

Suppose we were walking deep in the woods and a gleaming in the distance catches our eye. We go to investigate and as we approach, we see a bronze statue.



We can see that a certain quantity of bronze was used. The bronze is the material from which the statue was made. Aristotle considers this the material cause of the statue. The bronze is responsible for the statue in the sense that it is that from which the statue was made. Aristotle introduces the material cause:

"One way in which the word 'cause' is used is for that from which [τὸ ἐξ οὗ] a thing is made and continues to be made [γίνεται τι ἐνυπάρχοντος<sup>39</sup>]<sup>39</sup>—for example, the bronze of a statue, the silver of a bowl, and the genera of which bronze and silver are species" (*Physics* 2.3 194b24-27; translation Waterfield's).

The material cause is the stuff from which something is made. Other examples might include the flesh and bones of living things. The material cause can be specific: this bronze of this statue. Or more generic: there are many kinds of metal--bronze, silver, gold, etc. Depending on the why question, the specification of the material cause can be more specific or more general.<sup>40</sup> For example: why is this statue here? Because this quantity of bronze was available to be turned into a statue. This might be a question/answer format we would be happy with given that we found this particular statue here. But someone might want to know about why there are statues in the first place. And one answer citing a material cause could be: Because there are certain metals available. "Metal" is more of a genus, while bronze is a "species."

We do not yet have a complete account of the statue. The quantity of bronze could have been a number of things: some bullets (maybe?), a tea set, a table, or simply

---

<sup>39</sup> "Continues to be made" is not the best translation of ἐνυπάρχοντος. This is discussed more later on.

<sup>40</sup> Aristotle seems sensitive to this point here:

"Again, the general is related to the general, and the individual to the individual: a statue is produced by a sculptor, but this individual statue by such and such an individual sculptor" (*Physics* 195b27-29) (Translation Wicksteed and Cornford's (1957)).

left to be a quantity of unshaped bronze. What accounts for this being a statue instead of something else? I turn now to the formal cause.

### 3.2.2 The Formal Cause

As our investigation of the statue continues, we wonder about the shape of the statue. The features of Aristotle are portrayed in the way the bronze is structured. But as we noticed in the last section, the quantity of bronze could have been any number of things. Instead, in this instance, the bronze has been shaped into the dimensions typically attributed to Aristotle. The bronze has been shaped as to give an idea of his hair, his eyes, nose, mouth, and the rest of his body. The bronze was formed into the dimensions of Aristotle. This structure or shape (in this simple example) is the formal cause:

“A second way in which the word [cause] is used is for the form or pattern [τὸ εἶδος καὶ τὸ παράδειγμα] (i.e. the formula for what a thing is [ὁ λόγος ὁ τοῦ τί ἦν εἶναι], both specifically and generically, and the terms which play a part in the formula). For example, the ratio 2:1, and number in general, cause the octave” (*Phys.* 2.3 194b27-29; translation Waterfield’s).

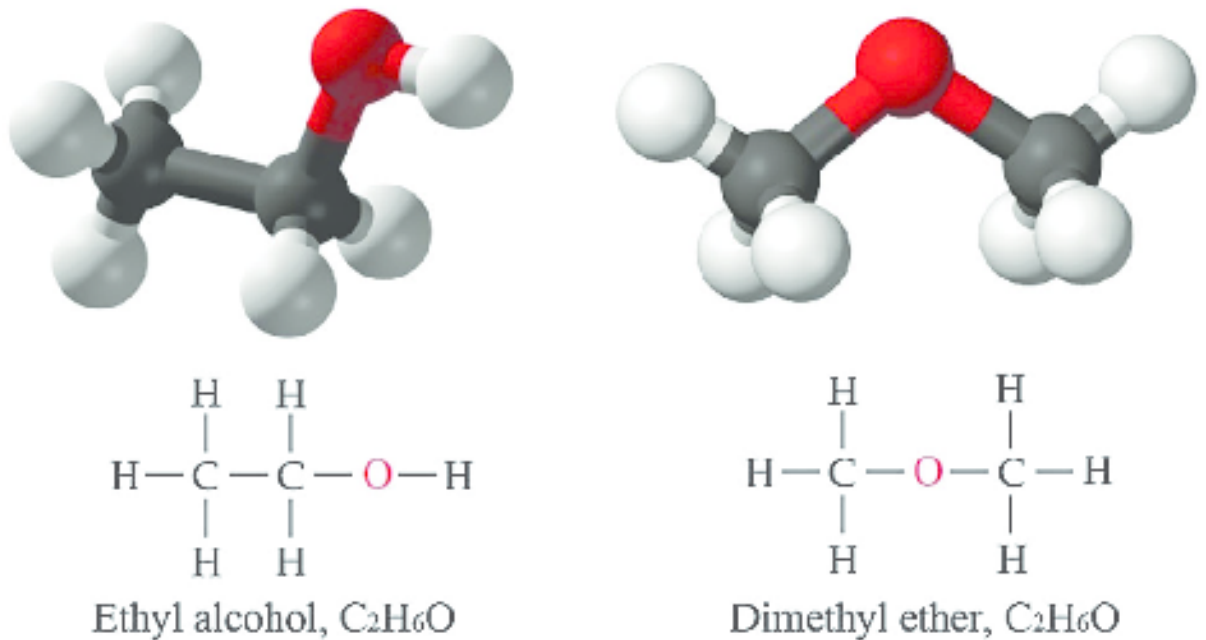
There are two things the formal cause explains:

- (1) the formal cause explains the structure of a particular thing
- (2) the formal cause explains why something is numerically one thing.

I discuss each in turn.

The εἶδος and the παράδειγμα account for the way a thing is structured or arranged. The account, or formula (ὁ λόγος), of this arrangement is what is given in an explanation. The λόγος gives the account of what-it-is-to-be (ὁ τοῦ τί ἦν εἶναι) something. Some examples from chemistry help us get traction with what Aristotle has in view here.

Ethanol and Dimethyl Ether are different things in the world, although they have the same number of carbon, hydrogen, and oxygen atoms. Both ethanol and dimethyl ether can both be described as  $C_2H_6O$ ; each has two carbon atoms, six hydrogen atoms, and one oxygen atom. Depending on the way these atoms are arranged--or structured--the subsequent substance has different properties. Here are depictions of two chemicals--Ethanol and Dimethyl Ether<sup>41</sup>:



If you structure these atoms one way, you get the stuff found in alcoholic beverages (Ethanol). Structured another way, you get a poisonous gas whose uses include being used as a fuel or refrigerant (Dimethyl Ether).<sup>42</sup> It is not enough therefore to simply have a pile of carbon, oxygen, and helium atoms because depending on how they are

<sup>41</sup> The figure below is from (Tumay, 2016).

<sup>42</sup> Of course, there are other ways of writing chemical formulas that express the structural difference. For example, the formula of ethanol can be expressed:  $CH_3CH_2OH$ . The chemical formula of dimethyl ether can be expressed:  $CH_3OCH_3$ . I used the depictions because they help bring home the important structural dissimilarities between them.

arranged, they have different properties. Thus, the form explains the structure or arrangement of a given thing.

Another point is in order. Ethanol and Dimethyl Ether have a form or pattern [τὸ εἶδος καὶ τὸ παράδειγμα]. But, the chemical formula  $C_2H_6O$  and the above images give the λόγος, the account, of what-it-is-to-be [ὁ τοῦ τί ἦν εἶναι] an Ethanol or Dimethyl Ether molecule, which essentially includes their structure (the depictions are more accurate than  $C_2H_6O$  because they include the structure) as well as the materials they are structured in. The εἶδος καὶ τὸ παράδειγμα are what is found in nature, whereas the λόγος is what features in explanation. And the λόγος essentially features the structure of the thing.

What the formal cause also explains is why this is one Ethanol molecule, why that is one Dimethyl Ether molecule, and why that over there is just a congeries of carbon, hydrogen, and oxygen atoms. That is, the congeries of atoms is 7 things (atoms) while the molecule of Ethanol is one thing: a molecule of Ethanol.

A question immediately arises: how does one tell the difference between a molecule of Ethanol and a congeries of atoms? Or the difference between a congeries of wood, nails, mortar, and a house? The answer is that, often but not always, the way something is structured confers powers the parts did not have on their own. For example, when the wood, nails, etc. are arranged in such and so a way, they can offer protection from the elements. When the two carbon atoms, the six hydrogen atoms, and the one oxygen atom are arranged in such and so a way, they have the further property of intoxication; when arranged in a different way, they have the property of being poisonous to humans.<sup>43</sup>

---

<sup>43</sup> The final cause also explains why some things are numerically one and not many: a particular thing can be unified by function. Functions can be the result of agency (hammers, etc.). Or they can be the result of biological processes (living organized systems and their parts).

The bronze statue of Aristotle can be treated in a similar way. The statue, like the molecules, is the result of certain material being shaped in a particular way. Also like the molecules, the statue (arguably) has several parts (including Aristotle's eyes, hair, mouth, etc.), but they are arranged such that they form a unity, a singular thing. Unlike the molecule, however, the shape of Aristotle could have been instantiated in a wider variety of materials. But even here, there are still material requirements: Aristotle's dimensions have to be instantiated in a solid that can hold the shape. This leaves out a number of materials, such as gases and liquids.

In short, we can answer two kinds of why questions by appeal to the formal cause: why is there a statue here? Because bronze was arranged into the shape of Aristotle (species of metal). Why are there statues? Because certain quantities of metal are available to be structured into shapes.

Another reason I selected the chemical example above is because the form also specifies the sort of materials required for a given thing. In the case of the molecules above, specifically these kinds of atoms are arranged. But even in the case of a house, the craft of housebuilding and what it takes to be a house implies certain materials (wood, nails, etc.) and not others (gin, socks, etc.). What explains this close connection between the formal and material causes is just that instantiated structures are ways suitable material is arranged. For this reason, Aristotle thought there were no uninstantiated forms since something's form is nothing over and above the way certain materials are arranged. But neither is there any un-formed matter. For matter itself is a something, and every something has form. Hylomorphism is the claim that everything has matter and form.<sup>44</sup>

---

<sup>44</sup> I am claiming on behalf of Aristotle that everything that exists has matter and form. But Aristotle himself reserved this claim for things in the sublunary world, but not the celestial realm. Moreover, his discussion of the Unmoved Mover seems to imply that the Unmoved Mover has form but no matter. Still, for the scope of this project, which is rooted firmly in the sublunary, these concerns can be set to one side.

Aristotle thinks this holds even at the foundational level of the world. Aristotle held that at there were only four primary contraries<sup>45</sup>: heat, cold, wetness, and dryness. These four contraries are not reducible to anything else:

"It is clear, then, that all the other differences are reduced to the first four, and these cannot be further reduced to a lesser number; for the hot is not that which is essentially moist or essentially dry, nor is the moist essentially hot or essentially cold, nor do the cold and the dry fall in the category of one another nor in that of the hot and moist; hence these must necessarily be four of these elementary qualities" (*On Coming-To-Be and Passing Away* 330a24-29) (Translation Forster's (1955)).<sup>46</sup>

But even these fundamental contraries are not found in isolation. Anna Marmodoro (2014), discussing Aristotle, says that they "...pair up and constitute the *four simple elements*: namely fire, air, water, and earth" (p. 7).

"Fire is hot and dry, whereas Air is hot and moist...and Water is cold and wet, while Earth is cold and dry (*On Coming-To-Be* 330b3-5) (translation Marmodoro's (2014, p. 7)).<sup>47</sup>

And these elements are always changing into one another as the constituent contraries act on one another:

"For these bodies [Fire, Water and the like] change into one another (they are not immutable as Empedocles and other thinkers assert, since alteration would then have been impossible), whereas the contraries do

<sup>45</sup> The term "contraries" is Anna Marmodoro's (2014).

<sup>46</sup> Δήλον τοίνυν ὅτι πᾶσαι αἱ ἄλλαι διαφοραὶ ἀνάγονται εἰς τὰς πρώτας τέτταρας. Αὗται δὲ οὐκέτι εἰς ἐλάττους· οὔτε γὰρ τὸ θερμὸν ὅπερ ὑγρὸν ἢ ὅπερ ξηρὸν, οὔτε τὸ ὑγρὸν ὅπερ θερμὸν ἢ ὅπερ ψυχρὸν, οὔτε τὸ ψυχρὸν καὶ τὸ ξηρὸν οὔθ' ὑπ' ἀλλήλ' οὔθ' ὑπὸ τὸ θερμὸν καὶ τὸ ὑγρὸν εἰσὶν ὥστ' ἀνάγκη τέτταρας εἶναι ταύτας.

<sup>47</sup> τὸ μὲν γὰρ πῦρ θερμὸν καὶ ξηρὸν, ὁ δ' ἀήρ θερμὸν καὶ ὑγρὸν...τὸ δ' ὕδωρ ψυχρὸν καὶ ὑγρὸν, ἡ δὲ γῆ ψυχρὸν καὶ ξηρὸν.

not change" (*On Coming-To-Be* 329a35-b3) (Translation Marmodoro's (2014, p. 8)).<sup>48</sup>

The upshot is that even the four simple elements have matter and form. The form is the way the contraries are related, while the contraries are that in which the form is instantiated.

Everything is composed of the elements but structured in different and often more complex ways; there are embedded hierarchies of hylomorphic entities (substances in substances). Take for example the human body. The whole human body can be thought of as a biological system. The material cause is the biological matter while the formal cause is the way this matter is structured. But the human body can be subdivided into more systems: the reproductive system, the endocrine system, etc. Each of these will have their relevant formal and material causes. Further each of these systems can be further subdivided into cellular systems. And these can be further subdivided...all the way down to the level of the four elements. As discussed in the material cause section, explanations can be more or less specific, focusing on broader unities, to more specific ones.

Another important upshot to this point is that Aristotle is sensitive to the embedded hierarchies of hylomorphism in explanation. For example, when I ask why is there a house here? I do not have to include in my answer the hylomorphic structure of each particular component all the way down to the fundamental elements. The relevant level of hylomorphic specificity is fixed by the question.<sup>49</sup>

The last point I want to make is that not only things, but their properties have matter and form. For example, Morgan has matter and form, but her psychological states are also hylomorphic. For example, as we shall see, Morgan (a hylomorphic

---

<sup>48</sup> ταῦτα μὲν γὰρ μεταβάλλει εἰς ἄλληλα, καὶ οὐχ ὡς Ἐμπεδοκλῆς καὶ ἕτεροι λέγουσιν (οὐδὲ γὰρ ἂν ἦν ἀλλοίωσις), αἱ δ' ἐναντιώσεις οὐ μεταβάλλουσιν.

<sup>49</sup> For a similar analysis of explanation and hylomorphism, see Shields (2014, p. 54).

substance) has hylomorphic properties. For example, suppose Morgan is angry. Her angry state is a structured phenomenon as well. This will be discussed more later on.

In summary, Aristotle's hylomorphism holds that everything has matter and form, even the fundamental elements of the world. Moreover, there can be embedded hierarchies of hylomorphic entities. Finally, the properties of a substance are hylomorphic.

The closeness of matter and form is also important later. Someone might argue that if we just account for the locations of all the bits of material in a statue, we get the form for free. That is, the material cause would obviate the need for the formal cause. However, this is not the case. But for now, we should continue our causal investigation of the statue. Up next is the efficient cause.

### 3.2.3 The Efficient Cause

We continue our investigation of the bronze statue. Our investigation has yielded the results that there is a material cause--the bronze is what the statue is formed in. There is also a formal cause--the bronze has been structured in a particular way. The structure partly explains why this quantity of bronze became a statue of Aristotle and not a coffee table; the forms are different. The shape of Aristotle was instantiated in the bronze, and not the form of a table. But this does not explain why fully yet.

How was this form instantiated in this bronze? Surely, the form did not spontaneously assemble. There is nothing in the nature of bronze to cause such a thing. What did? The answer to this question is the efficient cause. In this case, a likely candidate of efficient cause is the statue-maker. There are many steps to making a bronze statue and the statue-maker set them each in motion.<sup>50</sup> The reason the form of

---

<sup>50</sup> Actually, Aristotle would cite the craft of statue-making, or statuary, as the more accurate specification of the efficient cause. It is the craft that contains all the instructions; the statue-maker merely employs this body of knowledge. See: Falcon (2015). See also: *Metaphysics* 1013b6-9; *Phys.* 195a6-8; *Phys.* 195b21-25.



Aristotle was instantiated in this quantity of bronze is because the statue-maker set in motion the steps that would result in such a statue.

Aristotle introduces the efficient cause this way:

"Further, the primary source of the change and rest [ἔτι ὅθεν ἡ ἀρχὴ τῆς μεταβολῆς ἢ πρώτη ἢ τῆς ἡρεμύσεως] is spoken of as a cause, e.g. the man who deliberated is a cause, the father is the cause of the child, and generally the maker is the cause of what is made and what brings about change is a cause of what is changed [οἶον ὁ βουλευσας αἴτιος, καὶ ὁ πατήρ τοῦ τέκνου, καὶ ὅλως τὸ ποιοῦν τοῦ ποιουμένου καὶ τὸ μεταβάλλον τοῦ μεταβαλλομένου] (*Phys.* 2.3 194b29-33) (translation Shields' ).

The efficient cause explains the primary source of change [τῆς μεταβολῆς] and rest [τῆς ἡρεμύσεως]. There is another term that Aristotle uses elsewhere to indicate change—κίνησις. This latter term is also sometimes translated as motion. But whether motion or change, no one English word quite captures the scope of these terms in the Greek.

Christopher Shields (2007), commenting on these words, says:

“Most often it [κίνησις] means something wider than ‘motion’, if we restrict motion to local motion, or change through space...For Aristotle, [κίνησις] is much broader, encompassing also growth, diminution and simple alteration (*Phys.* 225a26-32, 243a6, *Met.* 1068.10; *DA* 406a1-15). Moreover, Aristotle has a second term, [μεταβολή], which is also best rendered as ‘change’, used sometimes interchangeably with [κίνησις] (*Phys.* 200b12-13, 218b20), though elsewhere distinguished from it (*Phys.* 225a34-b3, *Met.* 991a11)” (p. 434).<sup>51</sup>

<sup>51</sup> The omitted portion at the ellipsis is worth a mention: “(Although we may also speak of there being movement in someone’s bargaining position or in an economy or a market, we do tend in English to use ‘motion’ most centrally in the case of local motion.)”

Here are some examples to demonstrate the scope of κίνησις: the growth and decay of living organisms—plant, animal, and human; the ocean tides; the water cycle; the formation and healing of a pimple; the transformation that occurs when a cake is baked; the breeze passing by that caresses the face on a hot August day; the coming to be and passing away of anger, or fear, or any of the affections of the soul; an instance of instruction and learning; the behavior of the four primary elements. Change and often motion is intended to capture the whole world of change, alteration, and movement.

The efficient cause can also be the primary source of rest. Rest, like motion or change, has a much wider scope than just simply the cessation of local motion, or the cessation of movement through space. Rest also encompasses when growth stops, or when diminution stops, or when alteration stops. Rest refers to what stops the pimple from growing, when the learning/instruction event ends, or when an instance of anger stops.

Finally, an efficient cause can be a skill or a craft.<sup>52</sup> For example, in making a statue, there are at least two candidates for the efficient cause: the craft of statuary, the particular artisan who possess the craft of statuary. Which is the efficient cause? Both. Now, Aristotle thinks the most accurate specification of the efficient cause is the craft of statuary, but that is because the craft of statuary explains the processes that get set in motion to build statues. The particular artisan is incidental; it could have been him, her, or someone else. However, the artisan also sets in motion the processes and so too is an efficient cause. Moreover, there are contexts that require mention of the artisan as efficient cause, such as when we want to know why this particular statue looks this particular way, when other statues of the same type look a bit different.

We now have the material, formal, and efficient causes explained, but it still seems like there is something left out of the explanation of why there is a bronze of

---

<sup>52</sup> See: *Metaphysics* 1013b6-9; *Phys.* 195a6-8; *Phys.* 195b21-25.

statue of Aristotle. Why is the statue here on the hill in plain view of the valley? Why was bronze used for the statue instead of marble or some other metal?

### 3.2.4 The Final Cause--The "End" of the "Matter"

We have learned that the efficient cause set in motion all the relevant processes, but we do not know the answers to these questions:

1. Why were these processes set in motion now (and not later or sooner),
2. Why was the statue placed here instead of there, or indeed
3. Why was the statue made in the first place?

Answering these questions requires the reasons used by the statue-maker. The statue-maker might answer those three questions thus:

1. Because this is the only time I could get around to it,
2. Because I thought more people would be able to see the statue better,
3. Because I wanted to honor one of the greatest philosophers that has ever graced this green earth.

In the case of things like statue building and other such artefactual cases (where an intelligent builder employs his or her respective craft), the final cause involves the reasons for setting a process in motion. In this case, I would think 3. is the primary citation of the final cause when answering our original question: what caused this statue? The reason was to honor Aristotle, which the statue-maker did when she could and in full visibility of the public.

Aristotle introduces the final (or teleological) cause thus:

"Further, the end [τέλος] is spoken of as a cause. This is that for the sake of which [οὗ ἕνεκα] a thing is done, e.g. health is the cause of walking about. 'Why is he walking?' We say: 'To be healthy' — and, having said

that, we think we have indicated the cause” (*Phys.* 2.3 194b31-35, translation Shields’).<sup>53</sup>

As we saw, a final cause plays a role in making artefacts. But it also plays a role in human action. In Aristotle’s example, why is the man walking? The reason the man has for walking is to be healthy.

### 3.2.5 The four causes at work in explaining change

Let us review the four causes in the production of the bronze statue of Aristotle.

(Q) Why was the statue made?

(A) Because a statue-maker, wanting to honor Aristotle (final cause), set in motion a series of processes (efficient cause) that resulted in Aristotle’s shape (formal cause) to be instantiated in a quantity of bronze (material cause).

The four causes answer different aspects of the why question above. The statue was made:

- (a) Because suitable matter was available to be transformed into a statue (material cause).
- (b) Because the bronze was structured into the shape of Aristotle (formal cause).
- (c) Because the craft of statuary set in motion a series of processes leading to the statue (efficient cause).
- (d) Because the artisan wanted to honor Aristotle (final cause).

Aristotle would consider this a complete explanation and answer to the question.

---

<sup>53</sup> Aristotle also has an account of final causes that includes non-deliberated actions of living things. For example, the roots of trees stop growing when they locate a water source; spiders spin webs for the sake of catching food; certain trees grow leaves to protect their fruit. While I think Aristotle’s account of teleology in nature is very interesting and could fruitfully be discussed in the context of biological systems, it is beyond the scope of this dissertation. Since this is a work about psychological states, these are squarely in agential territory. However, for more information, see: *Physics* 2.8 and Christopher Shields (2014) discussion of Aristotle’s use of the final cause (pp. 91-104).

Let us now spell out the role of the four causes in explaining change. Recall that when the causes were introduced, Aristotle's aim was to explain change in nature:

"Obviously, then, this is what we have to do in the case of coming to be and ceasing to be, and natural change [τῆς φυσικῆς μεταβολῆς] in general. Then, once we know the principles [τὰς ἀρχὰς] of these things [of coming to be, ceasing to be, and natural change], we can try to analyse anything we are looking into in terms of these principles."

These are the four principles of change Aristotle gives us that he thinks account for all natural changes. Some of these changes can be fairly complex with many steps. For example, making a statue or constructing a house have many steps. In the case of making the house, the builder has to cut wood, bang in nails, pick up stones, shape and sweeten stones, and a whole host of other changes great and small. Aristotle thinks any of these changes will fit the four-causal schema. What I would like to do in what follows is first to explain the role each cause plays in change. Following this, I discuss how Aristotle thinks the causes can feature in explanation. Clearly one need not mention every single cause and step of building a house in order to answer the question: why is a house there?

### *3.2.5.1 Material Cause in Change*

The material cause is the subject of change (that which is changed). In the case of the house coming to be, there are a variety of materials including nails, wood, mortar, stone, etc.<sup>54</sup> Each of these undergoes a whole variety of different changes: the wood might be shortened through sawing (while the arm of the builder moves up and down...the arm also being a material that is undergoing changes, namely moving up and down); or the stone might move from here to over there, and eventually into the wall;

---

<sup>54</sup> And each of these are themselves hylomorphic structures. A nail is a certain amount and kind of metal that was changed into a nail.

the nails change from being here to hammered into wood, and so on. Eventually the result is that the materials come to constitute a house; all the changes in the process of house-building were directed toward this aim.

There are a whole variety of materials that undergo change. Even though there are many, many changes that occur, whenever change does occur there is some kind of material cause in play. Indeed, for change to occur in the first place, *something* has to undergo change. What all the changes have in common, then, is that some kind of material was altered or changed in some way. But what kind of changes occur? This is what the formal cause explains, and I turn now to that.

### 3.2.5.2 Formal Cause in Change

In the previous section, we saw how formal causes explain what it is to be something and that they are constituted in suitable material causes.<sup>55</sup> But they also feature in explaining a change from a *what* to another *what*. Formal causes explain what the material changes *from* and what it changes *into*. For example, in the case of a statue, the formal cause allows us to identify *this* as a quantity of bronze (it has certain ratio of copper and tin), which *turns into* a statue. This bronze is changed into a statue of Aristotle, with certain physical dimensions and features. Form in both cases allows us to understand what was changed, and what it was changed into.

Sometimes, many, many intermediate forms come to be and pass away on the way to a final product, such as in the case of building a house. If the reader will bear with me, I'm going to offer up a fairly wide range of changes that occur during the making of a house. The builder might take a certain length of lumber. The lumber itself is a hylomorphic substance that will become a part of the house. On the way to this end, it will undergo many, many changes. For example, at time  $t_1$ , it is the original length, width, and height. As the saw begins its work, at time  $t_2$ , the beam is changed because it

---

<sup>55</sup> Later I will discuss how the formal and material causes are related via a constitution relation.

is missing these bits of wood (now sawdust), at  $t_3$ , the form has changed since it lost more sawdust...and so on until  $t_n$  when a certain portion of the beam falls off. Indeed, the time slices between  $t_1$  and  $t_n$  could be counted even more finely grained and account for each individual speck of wood that is sawed off; or perhaps someone could count time slices from each shift in number of molecules or atoms as the saw does its work. Form accounts for all the changes, from major (going from lumber to a part of a house) to minor (the individual molecules of wood falling off).

We, of course, do not have a word for each micro-change (or even mid-level changes, such as when the lumber-is-half-way-cut-through). But the key point here is that many forms come to be (such as the form at  $t_{347}$ ) and pass away (such as when  $t_{346}$  came to be and passed away). And it is specifically form that accounts the changes from what to another what.

A similar treatment can be given of the builder's arm while sawing. At  $t_1$  his arm is down and his hand is gripped around the handle of the saw; at  $t_2$  his arm is raised, causing the saw teeth to cut through a quantity of wood, all the way until  $t_n$  when a section of the lumber falls off. As before, the time slices can be more or less finely-grained. And, if one so desired, one could be so finely-grained, that each time slice represents each particular neuron firing in the process of sawing. The point generalizes to any of the changes of form that occur during building the house.

There is a key observation to make: whenever such changes occur, they do so in virtue of changing the material causes. The initial dimensions of the wood are altered by working on the material cause, the wood itself. The changes in position of the arm occur through manipulating the neurons and muscular subsystems in play. Altering the form occurs by manipulating the matter. But this does not mean that the formal causes can be identified with the material causes. The wood could have become a statue of Hermes, or a table, rather than being a part of a house. Any particular instance of a thing is a

composite of matter and form. Or, for a more Aristotelian gloss, the wood is *potentially* a table or a part of a house, but *actually* becomes a table or part of the house when the materials constitute the form. The matter explains what is changed; the form explains what it is changed into.

There is another kind of change that occurs. Suppose the builder, during the course of his work, needs to move a stone from here to there. Or needs to move several boards of lumber from this location to that. In these sorts of case, the change does not involve the alteration of the lumber or the stones; rather they retain their original status and merely change location.<sup>56</sup> The builder is still bringing about a change, but he is not altering the subject of change in any way. Nevertheless, he effects this change in the usual way: he has to operate on matter. In these cases, he is not changing the form of the wood or stone. But he is able to move them in virtue of their material causes. That is, he moves the stone and lumber *qua* stone and wood.<sup>57</sup>

In chapter 4, I will be showing how the changes in form that occur by manipulating material causes help explain Aristotle's account of mental causation.

So how do changes occur? How does something come to be changed into something else? This is the job of the efficient cause, the primary source of motion and rest. I turn to this next.

### 3.2.5.3 *The Efficient Cause in Change*

The efficient cause is that which brings about the change. And, as discussed previously, there are at least two kinds of efficient cause that Aristotle specifies: the cause that sets in motion a change or series of changes (through some kind of contact)

---

<sup>56</sup> Of course, the builder's body changes quite a bit as he moves the materials: bending down, standing up, walking this way. In this case, the story given before is applicable.

<sup>57</sup> In the discussion of matter and form in these sections, I appear to be conflating two kinds of changes that Aristotle distinguishes between: substantial change and alteration. However, both kinds of change involve matter and form and that is all I need to make the ultimate point I am going to make about mental causation. Aristotle discusses substantial change and alteration here: *Phys.* 190b11-24.



and crafts or skills (such as statuary or housebuilding). I would like to bring this general structure to bear on the variety of changes that occur in examples like housebuilding.

At a general level the craft of housebuilding is the most accurate specification of the efficient cause of the house because it explains all the subsequent changes that occur that result in the finished product. But also, the particular builder that manipulates the materials into the finished product is an efficient cause.

There are also a series of smaller and micro changes that occur throughout the processes that result in the house. For example, when the board of lumber is being sawn through, this pull of the saw results in the change from the original state of the board at  $t_1$  to the change of it missing this tiny amount of wood at  $t_2$ ; that push of the saw results in the change from its state at  $t_2$  to having still less wood at  $t_3$ , and so on until the builder finishes his work. Each of these smaller changes to the lumber have an efficient cause: the push of the saw, or the pull of the saw. If we wanted to explain the micro-changes that occur, it would be more of the same: whatever the smallest measurement of the push or pull that results in this particular molecule being separated would be the efficient cause.

A similar story can be told about the builder who is sawing. The efficient cause of the saw being pulled or pushed will be a capacity he has to saw (capacities as efficient causes plays an important role in chapter 4). And at a neurobiological level, each firing of a neuron is the efficient cause of a subsequent firing of a neuron, which is the efficient cause of...the story will eventually encompass the movement of the relevant muscular subsystems.

Each step--macro and micro--of the way is determined by an efficient cause acting on an amount of matter. Each step of the way there is an efficient cause that serves as the source of motion or rest.

In the case of house building (and other similar processes), the craft of housebuilding plays an important role in specifying just why each step is taken and is, in this sense, present at even the microlevels of change.

Next, I turn to the final cause in change.

#### *3.2.5.4 The Final Cause in Change*

The final cause specifies the reasons for the changes that occur.<sup>58</sup> There is an end aimed at: a completed house, for example. It is for this end that the builder employs the knowledge of housebuilding and starts setting things in motion.<sup>59</sup> This cause explains why the course of action was set in motion in the first place.

Final causes can often be ranked hierarchically. If we look at the big picture of the building of the house, we can reference the aim the builder is holding: a finished house. But we might also wonder why the builder is sawing. The end of this goal is a board of lumber with the right dimension for a window, say. The overall goal is the completed house, but there could be a variety of subordinate aims under this: to get this board in the right dimensions for a window; to sweeten this stone for this part of the wall, etc. Such subordinate final causes are still linked to the ultimate final cause: the building of the house. It is just that several smaller processes must be engaged in.

Where this picture does not apply is at the smaller- and micro-levels. That is, each stage of micro- or smaller-level changes: neurons firing or this part of one push of the saw. At this level, it would be hard to come up with a hierarchically arranged sub-aim

---

<sup>58</sup> Before, in a footnote, I pointed out that Aristotle thought there were non-intentional final causes in nature: when a spider spins a web, or when trees throw down roots until they find a water supply, or when a lion grows teeth. Such things are not done consciously, but the biological systems in play have a clear end and it is hard to explain why the spider made the web, or why the roots stopped growing when they found water, or why the teeth are just so fashioned that they serve the lion's nutritive needs, without reference to the end. I am not in this example of housebuilding or later treating this very interesting expansion of the final cause. Since this project is ultimately concerned with mental causation, a reason or motive will suffice for the purposes here.

<sup>59</sup> The final cause is also part of the formal cause in states of desire. The builder's desire to build a house for the sake of earning a profit, or for the sake of having shelter, or some such. This will be discussed in more detail in chapter 4. For now, it is enough to specify the end as what the builder is aiming at.

that connects to the intention of making a board of lumber of the right size. But it doesn't need to either. Imagine the builder sawing the board of lumber. Although he does not have micro-aims for each stage of the movements of his arm movements, his arm movements are still in service of the aim of cutting the board (and eventually the finished house). We can still say that his arm is in motion for the sake of the board and this ranges over the variety of micro-movements that occur at any given point.

### 3.2.6 Taking Stock, and Some Loose Ends

In this chapter so far, we have seen what the four causes of Aristotle are supposed to be, and how they feature in change, especially changes that require a variety of different processes. But they are some loose ends that should be addressed. First, with so many instances of material, formal, efficient, and even final causes, how do we know which is relevant in any given explanation? Second, does Aristotle think at least one of every kind of cause should be cited in a given explanation? The short answer is no. Finally, what if there are more than 4 causes? I take these each in turn.

Which of the material, formal, efficient, and final causes is sufficient for an explanation? That answer is dependent on the explanatory interests of the questioner. For example, if I ask: why is there a house here? A complete answer would be because a builder wanted make a house (final house) and so set in motion (efficient cause) the processes that would lead to it by changing relevant materials into a house (material and formal causes). But an interested person could also ask why the builder cut this board of lumber. The answer would be to shape it to the correct dimension (final cause) and so started to saw (efficient cause) the board (a formal/material composite). The efficient cause also acts on the materials that constitute the forms. It is this explanatory interest point, I think, that was a major contributing factor in interpreting Aristotle as an

explanatory pragmatist. At any rate, the explanatory interests of the questioner will determine the correct answer to this question.<sup>60</sup>

Does Aristotle think that all four causes must always be cited in a given explanation? No. Aristotle does not claim that all four causes are operative in a given context. For example, rain storms lack a final cause (as do things like volcanic eruptions and water freezing into ice).

Lastly, what if there are more than four causes? While Aristotle thinks there are only the four causal principles, this does not mean there couldn't be a fifth or sixth or more. However, Aristotle issues something like a challenge about this<sup>61</sup>:

"We have investigated these [four causes] sufficiently in the *Physics*; however, let us avail ourselves of the evidence of those who have before us approached the investigation of reality and philosophized about Truth. For clearly they too recognize certain principles and causes, and so it will be of some assistance to our present inquiry if we study their teaching; *because we shall either discover some other kind of cause, or have more confidence in those which we have just described*" (*Metaph.* 983b1-7) (translation Tredennick's (1933) (italics mine).

Aristotle goes on to investigate whether or not his predecessors had indeed described another kind of cause and concludes they had not. But the subtle challenge could be reformulated to include not just predecessors, but anyone. If someone were to discover a fifth kind of cause, it would not hamper the present investigation. This is because as shall be shown in chapter 4, whatever the fifth kind of cause would turn out to be, the original four are clearly operative in cases of mental causation. An advocate of the fifth

---

<sup>60</sup> Jaworski thinks that different explanatory contexts reveal different causes and causal relations. The key difference between Jaworski and Aristotle on this point is that Aristotle would say that these causes and causal relations can be further subcategorized under the four causes. But, for a good assessment on the structure of how- and why-questions, see: Jaworski (2009).

<sup>61</sup> I am following Shields (2014, p. 53) in interpreting Aristotle's comment as a challenge.

cause would have to show whether or not the sort of cause was operative in a case of mental causation on the one hand, and, on the other, whether or not the cause takes up the explanatory resources of one of the original four. I am offering Aristotle's four causes as having a worked-out metaphysic as to how each of the causes is related to the next. This work, of course, would have to extend to any fifth or sixth cause. That is, how is fifth-cause related to each of the other four? But until such a time comes, I think Aristotle offers enough for us to get on with.

What remains to be explained is why these four causes are not in competition with one another, and how they even require one another.

### 3.3 Why the Four Causes do not Compete

The key ideas of this section are that there is good reason to think each of the four causal principles explain different features of the world, and so are not in competition with one another, and in fact even require one another.

If Aristotle is right, each cause has a distinctive job to do. The material cause is that which gets structured into something. The formal cause is supposed to explain why this quantity of bronze is a statue and not a bullet, trash bin, or some other artefact. The efficient cause is supposed to explain what set the process(es) in motion that resulted in the change, and the final cause, when present, is supposed to explain the reason why the process(es) were set in motion. If any of the causes "compete," it means that the work of one cause is actually done by another cause.

It seems clear that the efficient, final, and material causes have their own distinctive jobs to do. The efficient cause is the primary source of change or rest. The efficient cause *acts on* matter and *brings about* different, structured things. What is acted on and the primary actor do seem to be distinct from one another. As for the final cause, a reason for setting something in motion does seem different than what sets

something in motion, and what is acted on.<sup>62</sup> For example, returning to the baseball pitcher from earlier, throwing the ball with a reason (to strike out the batter) is different than the triggering mechanisms that bring this about. And similarly, for Penfield's experiments. The reason a patient's arm moved was because the brain was stimulated and not because she herself had a reason. And the material cause doesn't seem at risk of redundancy; any instance of change requires something that is changed from something to something. These particular atoms are changed into dimethyl ether or ethanol; they change from 6 atoms into one molecule of one or the other. The baseball pitcher's body changes from a standing position, to one of releasing the ball in such-and-so a way. This quantity of bronze is changed from the bronze into a statue.<sup>63</sup>

However, one might worry that the material cause nevertheless does the work of the formal cause, and so excludes it from the causal story. Suppose a physicist was looking at a bronze statue of Aristotle, and is able to specify the location of each fundamental particle in the statue, as well as all the relations to other fundamental particles each particle stood in. By accounting for all the matter of the statue, and the relations it stood in, it seems like form has no work to do; the shape is "thrown-in" for free.

But there are two jobs that can only be done by the formal cause: individuation (why this is an dimethyl ether molecule and not a molecule of ethanol, or just 6 atoms), and capturing commonalities between things made of different matter (how the dimensions of Aristotle, for example, could be constituted by different mediums such as bronze, paint, or marble). As for individuation, notice that it is not possible to neatly partition off the particles of the statue from the rest of the universe if we appeal just to

---

<sup>62</sup> Again, Aristotle argues for a kind of non-agential teleology in nature (e.g., trees grow leaves in order to protect fruit; but the tree is not consciously growing leaves, holding protection as *a reason*).

<sup>63</sup> This can work the other way; a statue can change back into bronze; the baseball pitcher can return from a shape of baseball throwing into standing again; and the chemical bonds holding the atoms together can be broken and the molecule of ethanol returns to 6 atoms.

the material cause. Consider; the fundamental particles of the statue are also standing in various relationships with particles "just outside" the region of the statue, and beyond. What cannot be accounted for merely by the material cause is why this particle is part of the statue, and that particle is an oxygen atom in the surrounding air.<sup>64</sup> If one wanted to insist that this particle is a particle of bronze, and that a particle of air, one has snuck in form through the backdoor. If bronze and air are different things, they are so ultimately because of the arrangement of their atoms. Bronze is a ratio of tin and copper, which in turn are composed of certain atoms. Configuration and structure explain why this is bronze and that is air if one wants to insist that the bronze of the statue and the oxygen of the air should be understood as distinct things.

Form is also what allows us to account for structures imposed across a variety of media, not just bronze; a painting, statue, and bust of Aristotle can be identified through use of form. The atomic elements of paint, bronze (or wood, or ice), or marble will all be arranged in various relations. Moreover, Aristotle might be in different aspects: in the marble bust, his face and part of his neck and shoulders are facing forward; in a painting, he might be depicted as walking with Plato discussing forms, and in a statue be sitting and thinking. Not only will accounting for all the particles in the painting not look anything like the arrangement of particles in the statue, there is no reason to think that the structures the relevant particles compose having anything to do with one another; the walking Aristotle is not shaped like the sitting one, for example. However, if we concede that Aristotle had a certain structure and shape that could be depicted or

---

<sup>64</sup> Of course, there are philosophers and scientists who think this is the case. In mereology, mereological nihilism is a thesis that states there are only fundamental properties in various relations; there are no higher order structures like ducks, chairs, and tacos. There are only particles arranged in various ways; our talk of objects is a relic of our cognitive processes. For more discussion on non-eliminative mereological nihilism, see: Gabriele Contessa (2014). For a different discussion on mereological nihilism, see: Jeffrey Grupp (2006). But I am presupposing for the sake of argument that such a radical, revisionary view is false.

imitated in various ways, then this explains why the painting, statue, and bust are all of Aristotle.

In short, the formal cause does some explanatory work that cannot be done by the material cause. So we can conclude that each of the four causes has its own explanatory work to do.

Moreover, it's not merely the case that the four causes aren't in competition. Indeed, each of the causes requires the others. To see this, consider again the kinds of why-questions that the four causes explain:

Why is this statue here?

- (1) Because a suitable quantity of bronze was available (material cause).
- (2) Because the shape and features of Aristotle were instantiated in the bronze (formal cause).
- (3) Because an artisan set in motion and completed a series of steps resulting in the statue (or, because the craft of statuary was employed) (efficient cause).
- (4) Because the artisan wanted to honor Aristotle (final cause).

Suppose the artisan wanted to honor Aristotle with a bronze statue, but there was no bronze to use. Or suppose he had the bronze but didn't know about the shapes or features of Aristotle. Or suppose the craft of statuary was not set in motion, even if the knowledge of Aristotle's features, bronze, and the intention to honor Aristotle were all present. Each of the causes, when operative, imply the others.

The general point is that an efficient cause requires something to act on and a final cause (when present) also requires an efficient cause and suitable materials, or the goal cannot be achieved.



### 3.4 How Matter and Form are Related

How are matter and form related according to Aristotle's metaphysics?

Answering this question is important because it will play a critical role in chapter 4 in understanding how a rationalizing cause and triggering causes are related. I proceed in the following way. First, I discuss two meanings of δύναμις--potentiality and power (or capacity)--in Aristotle's work that directly bear on the relation between matter and form. Next, I introduce a key passage where Aristotle discusses the relation between matter and form. Then, following Cohen (2016), I argue that Aristotle is best interpreted as proposing that the relation is what we might call *constitution*: the wood of a table constitutes the table; it is not *identical* with the table. The table (and house, etc.) are identical to the form/matter composite.

The heart and soul of Aristotle's account of matter and form is his distinction between potentiality (δύναμις) and actuality. δύναμις does double duty in *Metaphysics* Θ. δύναμις can be translated as "potentiality" or "potency" and even "power" or "capacity."<sup>65</sup> Aristotle says the primary meaning of δύναμις is the:

"...source of change in some other thing, or in the same thing *qua* other<sup>66</sup>" (*Metaphysics* Θ.1, 1046a1, translation Tredennick's (1933)).

This can be helpfully compared to Aristotle's definition of the efficient cause:

"the primary source of the change and rest [ἔστι ὅθεν ἡ ἀρχὴ τῆς μεταβολῆς ἢ πρώτη ἢ τῆς ἡρεμῆσεως] is spoken of as a cause...(translation Shield's (2014)).

Cohen (2016) notes that for Aristotle, the "exercise of such a power is a [κίνησις]--a movement or process." Thus, δύναμις in this sense is closely related to κίνησις and

<sup>65</sup> See Cohen (2016) and Anna Marmodoro (2014, p. 3).

<sup>66</sup> ...ἢ ἐστὶν ἀρχὴ μεταβολῆς ἐν ἄλλῳ ἢ ἢ ἄλλο. Also, an example of something acting on itself qua other would be a seed turning into a tree; the efficient cause is within the organism itself, acting on its relevant parts.

ἡρεμίᾳ, translated as rest. I set this sense of δύναμις to one side for now until the relevant sections. Now, I turn to the other sense of δύναμις Aristotle employs in *Metaphysics* Θ.

There is a second sense of δύναμις that Aristotle identifies that, as Cohen (2016) rightly notes, is better translated as "potentiality." In this second sense, Cohen (2016) says it is not "a thing's power to produce a change but rather its capacity to be in a different and more completed state." For example, a quantity of wood into a table or a statue, or a sapling has the capacity to develop into a more complete state, namely a full-grown tree. δύναμις in this sense, says Cohen (2016) is related to actuality, (ἐνέργεια).

Aristotle does not use any term that directly translates as "constitution", but the notion seems to be implicit in his discussion of actuality and potentiality:

"Since we have now dealt with the kind of potency [δύναμις] which is related to motion [κίνησις], let us now discuss actuality [ἐνέργεια]..." Actuality means the presence of the thing, not in the sense which we mean by "potentially." We say that a thing is present potentially as [a statue of<sup>67</sup>] Hermes is present in the wood, or the half-line in the whole, because it can be separated from it; and as we call even a man who is not studying "a scholar" if he is capable of studying. That which is present in the opposite sense to this is present actually. What we mean can be plainly seen in the particular cases by induction; we need not seek a definition for every term, but must comprehend the analogy: that as that which is actually building is to that which is capable of building, so is that which is awake to that which is asleep; and that

---

<sup>67</sup> This emendation is taken from Ross's 1941 translation of the same text. It can be found in the second volume of Jonathan Barnes's collection of Aristotle's works: Aristotle (1984).

which is seeing to that which has the eyes shut, but has the power of sight; and that which is differentiated out of matter to the matter; and the finished article to the raw material. Let actuality be defined by one member of this antithesis, and the potential by the other. But things are not all said to exist actually in the same sense, but only by analogy--as A is in B or to B, so is C in or to D; for the relation is either that of motion to potentiality, or that of substance to some particular matter" (*Metaphysics* Θ.6 1048a25-27...1048a32-b9) (translation Tredennick's, underlining mine).<sup>68</sup>

In this passage, Aristotle is discussing both substances (things that have matter and form) and capacities of substances (which also have matter and form). The example of a substance would be the statue of Hermes, and the example of a capacity would be a person's capacity for studying. The idea is that a chunk of wood can potentially be a statue of Hermes and becomes an actual statue of Hermes when wood is structured in this way. This corresponds to a thing's capacity to be in a different or more completed state (δύναμις) and actually being in the state (ἐνέργεια).

Similarly a person who is a scholar (and so has the capacity for engaging in scholarly activities), but not currently engaged in scholarship, is potentially a scholar and is actually engaging in scholarship when the capacity is activated (much like the person who has the capacity for vision is only potentially seeing when asleep, but is actually seeing when awake). These examples correspond to Aristotle's other sense of

---

<sup>68</sup> Ἐπει δὲ περὶ τῆς κατὰ κίνησιν λεγομένης δυνάμεως εἴρηται, περὶ ἐνεργείας διορίσωμεν... ἔστι δὴ ἐνέργεια τὸ ὑπάρχειν τὸ πρᾶγμα μὴ οὕτως ὥσπερ λέγομεν δυνάμει· λέγομεν δὲ δυνάμει οἷον ἐν τῷ ξύλῳ Ἑρμῆν καὶ ἐν τῇ ὅλῃ τὴν ἡμίσειαν, ὅτι ἀφαιρεθεὶς ἂν, καὶ ἐπιστήμονα καὶ τὸν μὴ θεωροῦντα, ἂν δυνατός ἢ θεωρῆσαι· τὸ δὲ ἐνέργεια. δῆλον δ' ἐπὶ τῶν καθ' ἕκαστα τῆ ἐπαγωγῆ ὁ βουλόμεθα λέγειν, καὶ οὐ δεῖ παντὸς ὄρον ζητεῖν ἀλλὰ καὶ τὸ ἀνάλογον συννοῶν, ὅτι ὡς τὸ οἰκοδομοῦν πρὸς τὸ οἰκοδομικόν, καὶ τὸ ἐγγηγορὸς πρὸς τὸ καθεῦδον, καὶ τὸ ὄρων πρὸς τὸ μῦον μὲν ὄψιν δὲ ἔχον, καὶ τὸ ἀποκεκριμένον ἐκ τῆς ὕλης πρὸς τὴν ὕλην, καὶ τὸ ἀπειργασμένον πρὸς τὸ ἀνέργαστον. ταύτης δὲ τῆς διαφορᾶς θατέρῳ μορίῳ ἔστω ἡ ἐνέργεια ἀφωρισμένη θατέρῳ δὲ τὸ δυνατόν. λέγεται δὲ ἐνέργεια οὐ πάντα ὁμοίως ἀλλ' ἢ τῷ ἀνάλογον, ὡς τοῦτο ἐν τούτῳ ἢ πρὸς τοῦτο, τόδ' ἐν τῷδε ἢ πρὸς τόδε· τὰ μὲν γὰρ ὡς κινήσεις πρὸς δυνάμιν τὰ δ' ὡς οὐσία πρὸς τινὰ ὕλην.

δύναμις in the sense of "power" or "capacity," that is contrasted with κίνησις and ἡρεμίᾳ (activity and rest). Both kinds of examples share an analogy with potentiality and actuality. The wood-to-table class of example are like non-existence to existence, while the vision and scholarship class of examples correspond to non-activation to being activated.

Cohen thinks that a contemporary philosopher wanting to make the same point could do so with the constitution relation between matter and form. Using the example of a table (which would correspond to the Hermes example), Cohen (2016) says:

"The idea here is that it is not the wood qua wood that is actually a table, but the wood *qua* table. Considered as matter, it remains only potentially the thing that it is the matter of. (A contemporary philosopher might make this point by refusing to identify the wood with the table, saying instead that the wood only constitutes the table and is not identical to the table it constitutes.)"

The wood of the table is not transformed into a table; one cannot say that the wood itself, as wood, just is the table. Rather the wood *constitutes* the table in virtue of being structured in that way. The wood constitutes the form of the table. The table itself, as actually being a table, is the composite of the form and matter. The table is identical to the form/matter composite. Actuality is tied to a certain quantity of material bearing a form. Cohen (2016) has a similar interpretation in view:

"His idea might be that not only can a piece of raw wood in the carpenter's workshop be considered a potential table (since it can be transformed into one), but the wood composing the completed table is also, in a sense, a potential table.<sup>69</sup>"

---

<sup>69</sup> Cohen (2016) admits that Aristotle might hold the view that the wood becomes the table: "The as yet uncarved wood is only potentially a table, and so it might seem that once it is carved the wood is actually a table. Perhaps this is what Aristotle means, but it is possible that he does not wish to consider the wood

The thought is that the wood in the table could constitute (and currently actually constitutes) a table, but it could constitute something else, say a statue of Hermes; the wood is potentially many things. It's worth repeating here that form is what accounts for change, from what to what. The wood undergoes a change (through an efficient cause) that results in another form/matter composite: a table.

A similar treatment can be given to capacities like having the capacity for scholarship or vision. These are a bit different though. In the case of a chunk of wood becoming a table or statue of Hermes, the wood has not received the form. In this case, a person already possesses the capacity, but the capacity is currently not activated. The change being explained here is that from non-activation to activation; it is potentially active and becomes active. They are similar, however (in Aristotle's terms, they can be understood by analogy), because there are still material causes that need to constitute the use of the capacities. Vision, for example, has certain material conditions (an eye, visual cortex, etc.). When a person is engaged in seeing, these are activated through a structured sequence of neural firings, etc. (to update Aristotle's biology). Similarly, the capacity to engage in scholarship will be a capacity of a certain part of the brain.<sup>70</sup> When activated, a certain sequence of neural firings (and perhaps other capacities) will constitute the process of engaging in scholarly pursuits.<sup>71</sup> There will be more said on

---

to be a table." I prefer the interpretation that Aristotle has in view the constitution relation. I think it makes better sense of the distinction between matter and form, and better grounds Aristotle's explanation of change. At any rate, the main goal of this project is to offer a plausible interpretation of Aristotle and not settle exegetical debates.

<sup>70</sup> There is another reading of this text that does not follow the pattern Aristotle is using here. Suppose there is a person who is potentially a scholar in the sense that they lack the capacity, but could become one through the relevant learning. This change would be more akin to wood coming to constitute a table. But it is not quite the same because it is not a substantial change. The change is from not having a certain capacity to having a certain capacity. On this reading, we have the first sense of δύναμις and ἐνέργεια, but not in a substantial sense. This could be helpfully compared with Aristotle's discussion of teacher and student in *De Anima* II.5 417a21-b16.

<sup>71</sup> This is not to say that substances cannot also be the efficient cause of other changes. For example, a plant comes into being through the union of gametes. When they are more developed certain plants have the capacity to turn light into usable chemical energy.

this in Chapter 4, as certain capacities will be critical for understanding Aristotle's account of mental causation based on these causes.

There is one more point to make about the capacities like engaging in scholarship. When activated, they are also composites of matter and form. But they are also the efficient causes of subsequent action. Because the scholar is now engaged in scholarship this capacity (or cluster of capacities) can set in motion his arm and fingers to turn a page in a book, or to write down a few sentences on the subject. They can cause further change. That some capacities, themselves composites of form and matter, can in turn be the efficient cause of other changes will play a critical role in chapter 4.

In summary, the relation between matter and form is that of constitution. This can take two related forms. A certain quantity of material can come to constitute a substance such as when a chunk of wood comes to constitute a table or statue. The table or statue is identical to the form/matter composite. Alternately, a capacity can go from being inactive to active. Both kinds are related by analogy. Moreover, some capacities, when active, can be the efficient cause of further changes. I turn now to chapter 4 where these insights will be applied to philosophy of mind and mental causation.

## Chapter 4: How Aristotle's Causal Pluralism Solves the Exclusion Problems for Mental Causation

In the last chapter, Aristotle's four causes were discussed, showing that each has different work to do, and that each does not overlap the other causes. Finally, the metaphysical relations between the causes were treated. Chapter 2 ended with some difficult questions for Jaworski's use of rationalizing and triggering causes:

- (1) Why don't the rationalizing and triggering causes compete?
- (2) How are the rationalizing and triggering causes related?

In this chapter, I will address these questions through the Aristotelian framework offered.

This chapter proceeds in the following way. First, I look at how Aristotle defines psychological states in *De Anima* 1.1. Aristotle uses anger as his paradigm case, intending it to provide a template for definitions for other psychological states: happiness, sadness, etc. But what this dissertation requires is a case of mental causation: how a psychological state contributes to bringing about a subsequent event. For example, Morgan, desiring a beer, gets up from the couch. What is the relationship between her desire on the one hand, and the getting up on the other? How are the four causes at work in a case of mental causation? Following this, I explain why each of the causes do not compete and how they are related, mirroring chapter 3.

I then turn my attention to Jaworski's rationalizing and triggering causes. I argue that the rationalizing cause of an action is constituted in the triggering causes; they are a hylomorphic compound. Once this is established, I show how the rationalizing cause, in virtue of structure, plays a direct role in the kinds of events that come next, also constituted in relevant triggering causes.

#### 4.1 Aristotle's Four Causes and Mental Causation

In this section, I show how Aristotle defines psychological states, their coming to be and passing away, in terms of the four causes. I show where each cause is at work in his definition of anger. Following this discussion, I turn my attention to a case of mental causation. How does one psychological state contribute to a subsequent action? I shift the example to that of Morgan, who is thirsty, getting up from the couch to get a Pabst Blue Ribbon from the fridge. How does Morgan's desire for a beer contribute to her getting up? Answering this question will require introducing Aristotle's faculty of practical reason. Once explained, I show how the psychological state activates the faculty of practical reason, and how the faculty of practical reason sets in motion the next event; getting up. I show how each of the four causes is at work in such cases. Finally, I explain why the causes are not in competition with each other and how they are related; this will draw on the results of chapter 3.

##### 4.1.1 Aristotle's Account of Psychological States

The four causes answer different questions about how something came to be (or passed away). These same causes are applied by Aristotle to psychological states. Aristotle, seeking to define psychological states, uses anger as his paradigm example.<sup>72</sup> I start with Aristotle's general account of psychological states, and then look at it in a bit more detail, fielding some initial concerns. Here again is Aristotle's definition of anger from *De Anima* 1.1 403a25-27:

"...clearly the affections of the soul are formulae expressed in matter [τὰ πάθη λόγοι ἐνυλοὶ εἰσιν]. Their definitions must therefore be in harmony with this; for instance anger must be defined as a movement of a body,

---

<sup>72</sup> It is paradigm in the sense that he offers a definition of anger in *De Anima* that is a template for any psychological state.



or of a part or faculty of a body, in such a particular state roused by such a cause, with such an end in view" (translation Hett's (1995)).<sup>73</sup>

"Affections of the soul," for our purposes here, can be understood as "psychological states."<sup>74</sup> Even though Aristotle does not explicitly mention the four causes, he clearly has the four causes in view.

Aristotle says in his definition: "anger must be defined as a movement of a body, or of a part or faculty of a body [*κίνησις τις τοῦ τοιοῦδι σώματος ἢ μέρους*], in a particular state roused by such a cause." So anger is a kind of *κίνησις* of a body or a part or faculty of a body. According to the distinctions in place between potentiality and actuality, this suggests that psychological states are like potentially seeing and actually seeing when waking, or possessing a skill, say scholarship, which is inactive and becomes activated.<sup>75</sup> When the capacity is activated by a suitable efficient cause, the relevant body parts undergo a kind of process, which is the actualization of the state. Psychological states are enmattered capacities which are actualized when activated.

A psychological state is "...roused by such a cause" [*ὑπὸ τοῦδε*]. As discussed in chapter 3, Aristotle thinks there are only the four kinds of causes in play. The cause that fits the bill here is the efficient cause. What sort of efficient cause would be such that it

---

<sup>73</sup> "ἄλλοιον ὅτι τὰ πάθη λόγοι ἐνυλοὶ εἰσιν· ὥστε οἱ ὅροι τοιοῦτοι οἶον "τὸ ὀργίζεσθαι κίνησις τις τοῦ τοιοῦδι σώματος ἢ μέρους ἢ δυνάμεως ὑπὸ τοῦδε ἔνεκα τοῦδε"

<sup>74</sup> Christopher Shields (2016) also identifies the affections of the soul with psychological states when explaining in general terms what *De Anima* is about: "This contrasts with *De Anima*, which introduces as a question for consideration "whether all affections are common to what has the soul or whether there is some affection peculiar to the soul itself" (*De Anima* i 1, 402a3–5). That is, in *De Anima* Aristotle wants to know whether all psychological states are also material states of the body."

Certainly, the sort of psychological states I am interested in for this project fall under the umbrella of "affections of the soul." But the term differs in some contexts from "psychological states." For more information, see: H. N. Gardiner (1919), Fred D. Miller Jr. (1999), and especially Victor Caston (2006).

<sup>75</sup> It is different than the wood potentially being a table, but actually becoming one (although as Aristotle noted in the passage before, they are analogically similar).

would actualize an instance of anger in someone? It will likely be some kind of perceived injustice (*e.g.*, Morgan getting struck in the face for no reason by Barney).

The psychological state is directed toward a particular end: "...with such an end in view" [ἐνεκα τοῦδε]. The end, or aim, of anger is retaliation. When a person is in an angry state of the kind Aristotle has in mind, they desire retaliation of some kind. It is worth comparing how the final cause operates in the case of a house, and in the case of Morgan's anger. "Why is there a house?" The answer (final cause): "Because the builder wanted to provide a family with protection from the elements." (Or some such.) "Why is Morgan angry?" Because she desires revenge/retaliation. There is an important difference between these two answers. The house does not desire to protect someone from the elements. Rather, the design of the house was constrained by those who wanted to make something that protects others from the elements. In the anger example, Morgan herself, being angry, desires retaliation. The house, upon completion of building, satisfies the final cause of the builder. There is no work left to do. Morgan upon becoming angry is in a state of desire. She is angry and the state disposes her to retaliation. And should Morgan decide to pursue the end of anger, anger will help explain why certain actions are undertaken.

What is true for the definition of anger goes for the case of mental causation as well. Morgan, in a state of desire, chooses to get up and go to the fridge for an ice-cold Pabst. Her capacity for feeling thirsty is activated by a suitable efficient cause. Her capacity of practical reason is activated. Should Morgan decide that quenching thirst is a worthwhile goal at this time, her capacity of practical reason sets her body in motion in those structured ways that end in drinking the beer. The difference between the mental causation example and the psychological state example is only that there are more coordinated movements directed toward a goal.

Onward to the formal and material causes. "Anger must be defined as a movement of a body, of a part or faculty of a body...." [κίνησις τις τοῦ τοιουδὶ σώματος ἢ μέρους]. Anger is a kind of movement of a part of the body. Aristotle thought this was a particular way that blood and heat would surge around the heart. What we can infer from this is that the movement of the particular part of the body is a coordinated, structured event. It is not just blood and heat thrown any old which way around the heart, but a structured sequence of events. The formal cause can be identified with the dynamic structure, while the material causes are the particular materials that are structured. Certainly, Aristotle's biology requires updating. But the key insight is that suitable materials are available that have the capacity to be in a different or more completed state.<sup>76</sup> Given what we now know about correlations between psychological states and brain states, the most likely candidate for these materials are neurons. When there is a certain pattern of coordinated firing of a series of neurons, a certain psychological state is instantiated.

At this point we need to discuss Aristotle's turn of phrase: λόγοι ἔνυλοί εἰσι. Aristotle says that psychological states are "formulae expressed in matter"-- λόγοι ἔνυλοί εἰσι. What does this mean? ἔνυλοί is a compound word that is a combination of 'in' (Greek: ἐν) and 'matter' (Greek: ὕλη). So the term might also be translated as "en-mattered." As such, another suitable translation of the Greek phrase is: the affections of the soul are en-mattered formulae.

What does Aristotle mean by "formulae?" The Greek word is λόγοι, which is the plural of λόγος. We saw λόγος before when Aristotle was explaining the formal cause. The λόγος is what is given in an explanation and accounts for both the way suitable materials are arranged, and the kinds of suitable materials themselves.

---

<sup>76</sup> This was Cohen's (2016) point I used in Chapter 3.

Thus, Aristotle is claiming that psychological states are enmattered, they are constituted by matter. This might initially seem like an odd claim given Aristotle's hylomorphic commitments. Specifically, that any particular thing that exists is already a composite of matter and form. What makes psychological states unique in this way, and why coin a phrase about being enmattered to express this?<sup>77</sup> The point of contrast can be brought out with mathematics as an example, as Aristotle himself does in *De Anima* 1.1 403b17-19:

"We have said [or sought to say] that the affections of the soul are inseparable from the physical matter of living beings in the way in which anger and fear are inseparable and not in the way in which line and plane are" (translation Charles').

Charles (2009) says that:

"Lines and planes are separable from physical...matter in two ways: they do not require the existence of any specific types of such matter to exist (403b14-15) and they are grasped by successful abstraction from such matter (403b15)" (p. 5).

The thought is that even if we encounter triangles and squares in materials, we can abstract triangles and squares from said materials and reason about them without any concomitant errors in reasoning. Aristotle's point is that errors would occur in reasoning about psychological states apart from the matter they are actualized in. In this way, they are more like dimethyl ether and ethanol: one cannot abstract the shape of these without also the matter, or one will make mistakes about how they come about.

---

<sup>77</sup> Kostas Kalimtzis (2012) points out, "[w]e do not find this word anywhere else prior to its appearance in this passage and Aristotle used it only this one time as far as we can tell from the surviving corpus" (p. 80).

Care needs to be taken here. Reasoning about dimethyl ether and ethanol would indeed result in reasoning errors if their structure was abstracted from their matter.<sup>78</sup>

One cannot get an ethanol molecule if one takes 6 other kinds of atoms and bonds them in the same way. Chemistry examples are rife with λόγοι ἔνυλοί εισιν. But things are not so clear cut in other instances. For example, is the shape of Aristotle so attached to matter? One can approximate his dimensions in paint, marble, bronze, or even popcorn if one were so inclined. It would seem that the mathematical dimensions of Aristotle are more like line and plane and can be abstracted from the matter in which they occur.

The question for this project is: are psychological states more like ethanol or more like Aristotle's dimensions?<sup>79</sup> Aristotle is clearly claiming they are more like the former. But one might think that multiple realizability challenges this claim: if one defines anger by its functional role<sup>80</sup>, for example, couldn't this same role be instantiated in very different material systems, such as brains, or positronic brains, or alien systems, or perhaps even angelic systems? And if that's the case, aren't psychological states more like line and plane and can be successfully abstracted from their realisers?<sup>81</sup>

There are at least two replies to this. First, we could say on behalf of Aristotle that it is just the case that psychological states must be constituted by certain suitable

---

<sup>78</sup> Aristotle's own examples of essentially enmattered entities or processes include weaving and snub noses. See: *De Anima* 408b1-13 and *Physics* 194a5-14.

<sup>79</sup> Charles (2009, 2012) is tempted to think that only a very particular bodily process can embody a psychological state according to Aristotle. Caston (2009) disagrees, and so do I.

<sup>80</sup> Elsewhere, I discuss why Aristotle differs from functionalism in terms of motivation; his definition of psychological states is closer to psychological content itself, rather than functional role. Nevertheless, the point I am discussing here remains the same. Even if one defines psychological states purely in terms of content (which Aristotle does not do), can't this content be realized in different systems? Couldn't particular psychological states like anger occur with slightly different material configurations? Perhaps this neuron and not that one fires this time, say. Or, if we expand this to David Lewis' (1983) Martian, couldn't the content of pain be constituted by "W-bladders inflating?" In this way, psychological states might be more akin to a house or the shape of Aristotle. While formulae for houses require some reference to material to correctly reason about them, there is more than one way to make a house, and more than one kind of materials that can fulfil the required role.

<sup>81</sup> Aristotle thinks it is possible that some kind of psychological events are not essentially enmattered, such as those involving active intellect. See: *De Anima* III.5 430a10-25.

material realizers. But this might have unwelcome consequences. For one thing, this would seem to rule out angels (and perhaps even God) as having psychological states since they must be constituted in suitable material systems, which, presumably, neither God nor angels would have. Also, it really could turn out as a matter of empirical fact that psychological states do not require a specific material realizer. I think what we should say is that while Aristotle may have thought psychological states were more like chemical formulae, we need not follow him there. But this does not damage Aristotle's essential insight.

The second reply to the question is that even if psychological states turn out to be more like line and plane, they are like any other entity in the natural world; any particular thing has a form constituted in matter. Moreover, humans and animals are the sorts of creatures that have psychological states constituted in appropriate biological materials. A key insight of Aristotle's hylomorphism is that forms are constituted by suitable materials, whether or not they can be successfully abstracted from those suitable materials and reasoned about without errors, and that many structures play a critical role in what happens in subsequent events (like in mental causation).<sup>82</sup>

At this point, it would be useful to address another concern for hylomorphism, if only to set it to one side for the purposes of this project. The concern applies both to Aristotle's account of matter and form and the updated biological account that includes neurons. The question is: why should this configuration of blood and heat around the heart, or this pattern of coordinated neural firings be identified with a desire for retaliation? If the form of anger is the desire for retaliation, why should these structures of biological materials embody this? There appears to be no necessity to these structures constituting this form.

---

<sup>82</sup> Aristotle seems to recognize a certain sense of multiple realizability. He thinks that many animals, which he knew were structurally different from humans, could be in states of pain or desire. This suggests he thought that different creatures with different physiologies could still constitute psychological states.

It is helpful to compare these cases with Aristotle's house example. There are a variety of materials that could constitute a house: wood, stone, brick, even ice to name a few. Whatever the material constitution, a house comes into being when the materials constitute a form that fulfills the end: protection from the elements. The end constrains the structural possibilities the materials can constitute. This would hold true across possible worlds too. If the aim at any world is protection from the elements, then the forms that wood or schm-ood, brick or schmr-ick, can constitute are constrained by the end of protection from the elements (or schm-elements). Furthermore, the end constrains the kinds of materials that can constitute such a form. For example, wood could work as a material, but carbon dioxide could not. Neither could any gas or (most) liquid(s).

But in philosophy of mind, why should we think that the end of anger--retaliation--places any kind of comparable constraints on the structure of biological materials on the one hand, or even on the kinds of materials that could be so structured on the other? Even at this world, Aristotle's proposed biology and our current updated biology result in very different materials and structures. Blood and heat around the heart, and the specific configuration thereof, is very different both in terms of structure and material than neurons firing and the configurations thereof. So much so, there is no obvious constraint the end places on structure or the materials constituting structure.<sup>83</sup>

While these are very interesting questions for any hylomorphist, I will have to set them to one side for the purposes of this project. This project is concerned with how Aristotle's account of the causes can be used to advance the project of explaining how psychological states are related to the nervous system on the one hand and play a role in causation on the other. This project can proceed with the premise that things in the

---

<sup>83</sup> Further, even at this world, we could suppose A.I. could become angry via a structured configuration of other kinds of materials.

natural world are a composite of matter and form, whether or not these things are more like ethanol or line and plane.

However, I think I can advance two reasons that can help with this. First, if we want a more naturalized account of psychological states, the most likely candidate for materials that constitute them is the body. That is, certain dynamic structures are constituted by certain bodily parts. In Aristotle's limited biology, he identified blood and heat around the heart. In our updated biology, this is a sequence of neural firings. Second, there is room in hylomorphism for a wider range of materials to constitute certain forms. For example, the materials and configurations of materials that can constitute a house are more numerous than the materials and configurations that can constitute dimethyl ether. The latter requires one specific configuration of very specific atoms. The former has a wider variety of structures and materials that can result in a house. The upshot is that anger (and other psychological states) might admit of a wider variety of materials and structures that can constitute it. With these considerations in view, I set this issue to one side.

Let us take stock. Why is Morgan angry? Or, what caused Morgan's anger? Each of the four causes answers the question differently:

- (A) Material Cause--Because suitable biological matter was available (neurobiological matter).
- (B) Formal Cause--Because suitable biological matter was constituted in a certain way.
- (C) Efficient Cause--Because Barney punched her
- (D) Final Cause--Because she desires revenge.

Before continuing to mental causation, I want to set up how I'll be discussing the form and matter and of psychological states. The form of a mental state of the kinds I am using (orectic states, or states of desire) is the desire for a particular end. This is



constituted by a certain sequence of neurons firing. Later I will map Jaworski's account of rationalizing and triggering causes onto this account of psychological states.

With this in place, I want to turn attention to a case of hylomorphic mental causation. This is an important consideration and development in Aristotle's thought. Typically, emphasis in the literature is given to explaining how the four causes figure in an instance of a particular mental state. What needs further development is how one hylomorphic psychological state is related to a subsequent action; a hylomorphic theory of mental causation.

#### 4.1.2 The Four Causes at Work in a Case of Mental Causation

In chapter two, I argued that denying Causal Closure is the way to solve the problem of mental causation. But, by doing that, it means that an account of psychological states, how they are related to bodily states, and how they are related to their effects is all the more pressing. I have been proposing that Aristotle's hylomorphic account of psychological states and causation is the way forward here. In chapter 3, I went over the metaphysics of Aristotle's four causes, including how they are related to one another, and why each cause does not do the work of the others. In the next couple of sub-sections, I am going to show how each of the causes is at work in cases of mental causation. Following that, I am going to show why each of the causes does not do the work of the others and how they are related.

Broadly speaking, cases of mental causation can fall into two categories. First, a psychological state can cause an event directly, without mediation by deliberation. For example, suppose Barney punches Morgan; Morgan immediately aiming for revenge punches him right back. Second, a psychological state can cause an event indirectly, mediated by deliberation. For example, Morgan is thirsty and sitting on the couch. Through deliberation, she decides to get up from the couch and get the Pabst Blue

Ribbon from the fridge.<sup>84</sup> I look at how the four causes are at work in both kinds of case.

#### *4.1.2.1. Mental causation without deliberation*

Imagine the following scenario. Barney, currently outside in the schoolyard after lunch, is furious and wants to take it out on someone. He sees Morgan, staring off into the distance with that philosophical intensity. He found his target. He does not know, however, that Morgan is also an adept and long-time student of mixed martial arts. Barney walks up to Morgan and punches her to vent his rage (this is something like his aim, the final cause). Morgan, caught up in the heat of the moment, immediately punches back. She does not deliberate but strikes immediately out of anger--a brilliant right cross that rearranges his nose and drops him. The efficient cause of her state of anger, of course, is Barney's unfair punch.

So Morgan is now in an angry state, brought about by Barney's unfair punch. She undergoes a change from being not-angry to angry. Her state of anger is constituted by the particular patterned neural sequences. What happens next--namely Morgan's punching of Barney--is an instance of mental causation and also needs to be explained via the four causes.<sup>85</sup> Here are the four causes in play at  $t_1$ , just before Morgan throws the punch:

- (1) The efficient cause: Morgan's anger (instantiated in neuronal firings and is the sort of thing that has the capacity to change hands into fists and can rearrange noses as successful cases of revenge).

---

<sup>84</sup> For Aristotle, many of these sorts of cases where a person acts in accordance with an aim, whether via mediation or not, are cases where desire of some kind of is at work

<sup>85</sup> A concern arises in this formulation: was Morgan the efficient cause through her faculty of anger, or was the efficient cause the state of anger itself? For the purposes of this section, not much hangs on these formulations. Without question, Morgan has a capacity for anger, and it is through this capacity that she acts. But it is also true that had she not been angry, she would not have punched Barney. We can focus on anger as the efficient cause because this state is what sets in motion the body next. But we can do so with the background knowledge that anger is a capacity of Morgan herself.

- (2) The material cause: the material Morgan's arm is made of, neurons (and later the flesh of Barney's nose).
- (3) The formal cause: the structured sequence of neural firings (that will later result in her arm in such and so a shape landing on Barney's nose).
- (4) The final cause: retaliation.

The change being explained is the thrown punch and rearranged nose at ( $t_2$ ). Here are how the four causes explain what happens next and how they are related to the four causes that were in play in bringing about Morgan's anger.

Morgan's anger just is the *efficient cause* of Morgan's punching Barney. This is what brings about the change from  $t_1$  to  $t_2$ . It is related to  $t_2$  as the primary source of change. This is not surprising in itself. As discussed in chapter 3, many hylomorphic entities are imbued with further capacities in virtue of their structure. The house has the capacity of protecting from the elements; the mousetrap has the capacity to catch mice; a human cell has the capacity to conduct cellular respiration, etc.<sup>86</sup> Anger, or (possibly) any psychological state, has a further power of causing motion (without deliberation in this case) and rearranging noses made out of suitable matter.<sup>87</sup> The other three causes involve certain mechanisms--biological and psychological--by which the efficient cause takes place.

So, the *material cause* of Morgan's anger (neurons) are (efficiently) causally connected to the *material causes* of Morgan punching Barney (the motor cortex, the muscular sub-systems, etc. that make up her arm). This makes sense; anger would not

---

<sup>86</sup> Of course, these capacities differ in a few ways. The capacity to protect from the elements is more of a passive capacity, while catching mice is more of a dynamic, active capacity. Anger would be more like the latter in this case.

<sup>87</sup> A few very important considerations. First, in this example, we are looking at a case of non-deliberative action. Morgan acts without deliberating first. It is perhaps possible that some actions require deliberation first. For example, if someone desires to get a college degree, this is likely something that cannot be done without deliberation. Second, should we think of each individual orectic state as a *sui generis* capacity, or is there a higher-order capacity of desire of which each orectic state is a species? I do not need to settle this question here; however this goes, if Aristotle is right, either will require constitution by relevant material in the nervous system.

have the capacity it does to bring about retaliation if it were not constituted by materials that were closely connected to locomotion. (We think this material is the stuff is neurons and connections to the motor cortex, Aristotle thought this was blood and heat in the heart region.)

The *formal cause* of Morgan's anger--desire for revenge--is the starting point of the change from non-retaliation,  $t_1$ , to  $t_2$  the fist-meets-nose-as-retaliation *form*. As always, formal causes explain the change from one state to another. And also as always, forms are constituted by suitable materials. What brings about the change from  $t_1$  to  $t_2$ ? The efficient cause, which is the angry state of  $t_1$ . But this does not occur in this case without the final cause.

The *final cause* is the same as the formal cause in this case--desire for retaliation. The final cause explains why the punch was set in motion--to bring about an instance of retaliation. As such, the caused event--the instance of retaliation--inherits in some way the formal cause of Morgan's anger--desire for revenge.

Between  $t_1$  and  $t_2$ , Morgan is angry, desiring retaliation. Her anger is constituted by a sequence of neuronal firings. Guided by the desire for retaliation, these neuronal firings can set in motion subsequent patterns of neural firings, which eventually explains the rising of her right arm, making the fist, pushing her arm out, and rearranging Barney's nose. There are a number of changes that occur between  $t_1$  and  $t_2$ , and a variety of forms come and go--the raising of the arm, the making of the fist, the series of movements outward, etc. There is a point of correspondence here between this and the fine-grained housebuilding example in the previous chapter. At  $t_1$  the builder begins work. At  $t_2$  the house is finished. Between these two times, a variety of changes take place: walking, picking up materials, sawing, cutting, shaping, etc. A variety of changes come to be and pass away, all guided by a suitable final and efficient cause which direct the actions. Many changes--forms constituted by suitable materials--come

to be and pass away. Similarly the change from the arm down by her side to slightly raised, to her hand being formed into a fist, to her arm thrusting outward in a series of movements--all of these are changes, and many forms come to be and pass away between  $t_1$  and  $t_2$ .

The general picture is that Morgan's anger--a desire for revenge (formal cause) constituted by a sequence of neurons firing (material cause)--sets in motion (as efficient cause) a series of changes that results in an instance of retaliation. Retaliation is the goal (final cause) of the series of changes.

Let's take a look at an instance of mental causation that involves deliberation.

#### *4.1.2.2 Mental causation with deliberation*

Let us now consider a case of mental causation that involves deliberation. Morgan is sitting on her couch playing video games. As time passes, a series of biological states informs Morgan that she needs to replenish fluids. Morgan is thirsty. In the refrigerator is a Pabst Blue Ribbon beer. Morgan desires that beer, and, after deliberating, she decides to get up and get the beer.

The event to explain is Morgan getting a beer and how her initial state of desire to quench thirst contributes to bringing this about. The order of how the four causes are related to the event is in a bit different order from the previous subsection--formal, material, final, and efficient. There is another key difference between this subsection and the previous. I mention at the beginning of the change to be explained she is in a state of desire to quench thirst. But I do not initially focus on the role this desire plays on setting her motion. I focus on the varying changes/forms her body goes through from the initial state of being on the couch to the end form, standing by the fridge drinking the beer. The reason for this will become apparent in the discussion of the efficient cause where deliberation will be explained.

The *formal causes* account for the changes of shapes taken on by the various parts of Morgan's body as she moves from the couch to the refrigerator. And, as always, the formal cause explains the beginning and endpoint of a change. Just before the series of changes begins, Morgan is sitting on the couch, her feet propped up on the coffee table holding her video game controller. Her initial form then, is this general shape and way of sitting.<sup>88</sup> The end change might be the way she is standing at the fridge, beer can in hand, drinking. In her initial state, she is in a hylomorphic state of desire, in this case she desires to quench her thirst.

Morgan will change shapes many times on her way to the fridge. At time  $t_1$ , she is sitting on the couch; at time  $t_2$ , she is standing up with her hands down; at  $t_3$  she is shaped with her left hand forward, her right hand back, and one foot in front of the other; and so on until  $t_x$ , when she is finished with the process. There will be many forms that will come to be and pass away in the process; many changes will occur. This is the same as the building of a house. Before the completion of the house, many forms come to be and pass away: the builder hammering at  $t_1, t_2, t_3$ , etc. (where the different times indicate different positions of the hammer and arm); the builder sawing at,  $t_1, t_2, t_3$ , etc. (where the different times indicate the different positions of the saw and arm). This is to be expected with hylomorphism wherein form is what explains a change from a prior state to a subsequent state.

The discussion here has been at the level of Morgan herself and the movements of her body as a whole. But the same hylomorphic treatment can be given down through the scientific domains. At a biological/chemical level, there are many neurons firing in sequence. This leads to movements of muscular subsystems, which have their series of processes sequenced to bring about movement. At a physical level, a number of atoms

---

<sup>88</sup> She might also have been in a form like being-on-her-back-with-her-feet-over-the-back-of-the-couch-and-her-head-in-an-upside-down-position. Or in a laying-on-her-side-propped-up-on-her-elbow sort of position.

are moved in sequence from here to there. As we traverse down the domains, there are an incredible amount of changes occurring each microsecond. These too can be accounted for through form accounting for the change from a prior state to a subsequent state. I turn now to how the material causes are related from the initial state to the end state.

The *material causes* are the materials Morgan's body is made of. These beginning and end forms are constituted by the materials of Morgan's body (and later the beer and beer can). The biological materials and beer can constitute these forms in the same way that the wood constitutes the table, or the building materials the house from chapter 3. At the highest levels, there are muscles, skin, etc. These are further composed of cells of differing functions, and these are set in motion via neural cells firing in sequence. At the lowest levels, we have the fundamental physical entities that constitute these higher-level materials. The series of changes that transpire are constituted by these materials.<sup>89</sup> Suitable efficient causes that can set these in motion needs to be found, but first, we should discuss the final cause.

The *final cause* in this case is a *reason*, namely, to quench thirst. A reason guides the series of processes, from the getting up, to each step, to the fridge opening, to the opening of the beer, to drinking it. The reason is to quench her thirst. The aim of the builder is to put together a house; this reason explains the specific changes set in motion and why they stopped at the end. The final cause explains the reason for the change of form at the beginning and the final resulting form at the end; the change from one thing to another thing. In the case of mental causation, the

---

<sup>89</sup> Of course, psychologists, biologists, and physicists will have different explanatory interests. The point is that Aristotle's four-causal schema can be applied to whichever domain...with qualification. While the physicist may be interested only in certain interactions at the fundamental physical level, she will have to take on board that the changes in locations of many atoms will be because of Morgan's desire. For example, her body, and the fundamental physical entities that constitute it, move from near the couch region to higher above it *because of* her decision to stand up and get a beer. Physics is not causally closed on this picture.

reason explains why Morgan got up, walked to the fridge, opened the beer can, and drank. And it explains why she stopped once the aim was achieved. She might now decide through practical reason whether or not she should return to the couch, just stand by the fridge for longer, or go outside. But the changes from laying on the couch to drinking the beer by the fridge are governed and guided by the reason. And this reason explains the movement upward from the couch and toward the fridge, as opposed to upward from the couch and upstairs, say.

The reason also explains the biological and physical movements to a large degree too. The *reason* that the muscular subsystems, the neurons, and the relevant atoms are in movement is because of Morgan's desire, which set everything in motion. This is part of the consequence of denying Causal Closure of physical given that Morgan's reason has consequences all the way down. The last thing to discuss is the efficient cause of Morgan's movement.

What is the efficient cause? Well, you might have thought that it's the desire for a beer. This would follow the pattern in the non-deliberative case where the desire itself has the capacity to move the person. But since this case involves deliberation about whether or not to satisfy this desire, matters aren't quite so straightforward. The question is what set Morgan in motion? Her desire to quench her thirst, or her decision/reason for getting up from the couch? The answer is both. Let's see how.

Aristotle himself recognizes that desires by themselves are not always sufficient for action.<sup>90</sup> Here is a key passage from *De Anima* 3.10 433a10-20:

"These two then, appetite [desire] and mind [ὄρεξις ἢ νοῦς], are clearly capable of causing movement if, that is, one regards imagination as some sort of thinking process...Both of these, then, mind and appetite, are productive of movement in space. But the mind in question is that which

---

<sup>90</sup> See: *De Anima* iii 9 433a6–8; cf. *Nicomachean Ethics* i 13, 1102b26.



makes its calculations with an end in view, that is, the practical mind [πρακτικοῦ νοῦ]...And every appetite is directed towards an end; for the thing at which appetite aims is the starting-point of the practical mind, and the last step of the practical mind is the beginning of the action. So these two, appetite and practical thought, seem reasonably considered as the producers of movement..." (translation Hett's).<sup>91</sup>

In discussing this passage, I want to leave to one side Aristotle's comments on imagination. I would like to focus on how appetite (or desire) and the practical mind work together.

Aristotle thinks that the practical mind (what others have translated as practical reason) and desire are able to produce movement. Morgan desires a drink to quench her thirst. Thirst quenching, as an end, is what practical reason picks up. Once picked up, the process of practical reason begins which ends in a decision to pursue the beer (at which point, it sets the body in motion), or to forego it for now (at which point, Morgan continues playing her video games). A question arises: how are desire and faculty of practical reason causally related?

I think a good way to understand this is as two efficient causes, each of which is bringing about changes from one state to another. In Morgan's initial state, she is happily sitting and playing video games, not in a state of thirst. Thirst is *brought about* by certain biological processes that cause her to change from a non-thirsty state, to a thirsty state. This is the efficient cause of the desire for a beer. In turn, this

---

<sup>91</sup> Φαίνεται δέ γε δύο ταῦτα κινουῦντα, ἡ ὄρεξις ἢ νοῦς, εἴ τις τὴν φαντασίαν τιθεῖν ὡς νόησιν τινα: ... ἄμφω ἄρα ταῦτα κινητικὰ κατὰ τόπον, νοῦς καὶ ὄρεξις, νοῦς δὲ ὁ ἕνεκά του λογιζόμενος καὶ ὁ πρακτικός· διαφέρει δὲ τοῦ θεωρητικοῦ τῷ τέλει. καὶ ἡ ὄρεξις (δ') ἕνεκά του πᾶσα· οὗ γὰρ ἡ ὄρεξις, αὕτη ἀρχὴ τοῦ πρακτικοῦ νοῦ, τὸ δ' ἔσχατον ἀρχὴ τῆς πράξεως, ὥστε εὐλόγως δύο ταῦτα φαίνεται τὰ κινουῦντα, ὄρεξις καὶ διάνοια πρακτική· τὸ ὄρεκτὸν γὰρ κινεῖ, καὶ διὰ τοῦτο ἡ διάνοια κινεῖ, ὅτι ἀρχὴ αὐτῆς ἐστὶ τὸ ὄρεκτὸν. (Aristotle & Hett, 1995)

Indeed, Aristotle thinks that most if not all animal locomotion, including human, is motivated by desire. Whether or not this is true is not the key focus of this project. Certainly, in the cases of mental causation that feature in the exclusion problem, desire typically plays a role.

desire is the efficient cause of her coming to reason about whether or not she should pursue her goal.

Shields (2016) seems to think along similar lines:

"...So, practical reason and desire act corporately as the sources of purposive motion in all animals, both human and non-human (*De Anima* iii 10, 433a9–16), even though, ultimately, it is desire whose objects *prick* practical intellect and *set it in motion* (*De Anima* iii 10, 433a17–2). For this reason, Aristotle concludes, there is a faculty of desire whose activities and objects are primarily, if not autonomously or discretely, responsible for *initiating end-directed motion* in animals.

What animals seek in action is some object of desire which is or seems to them to be good" (italics mine).

Desire is ultimately responsible for *pricking* practical intellect and *setting it in motion*. This suggests that desire acts as an efficient cause; the subject changes from a state of not engaging in practical reason, to a state of engaging in it. And then, once the subject decides through the faculty of practical reason to move, it sets the body in motion.<sup>92</sup> But it is the formation of orectic states that is responsible for setting the faculty of practical reason in motion.<sup>93</sup>

---

<sup>92</sup> It should be emphasized here that the specific details of how the various capacities in play interact not the main thing for this dissertation. For example, I am interpreting Aristotle as having two efficient causes that are often jointly sufficient for setting Morgan in motion. Morgan gets thirsty and desires to drink, which in turn results in the setting in motion of her practical reason. If she decides to pursue the end, the faculty of practical reason sets her body in motion. But one might also think that to perceive something as a possible good, *eo ipso*, sets in motion the faculty of practical reason. In this case, the perception of a possible good and the setting in motion of practical reason are simultaneous. Or again, the faculty of practical reason might be composed of any number of small cognitive abilities.

But, the more general points are these. (1) Whatever the in-depth biological story, it is the combination of orectic states and cognitive states at work in practical reason that functions as the efficient cause of voluntary action. (2) Whatever the in-depth biological story, if one faculty or capacity is setting in motion another--that is, causing a change which includes the structuring of biological materials--then we are dealing with efficient causes which bring about the change from one prior state to a subsequent one. Efficient causation is still in view.

<sup>93</sup> It might be argued that this makes it sound the formation of desire is a matter of direct bodily perception. Morgan becomes aware she is thirsty through her senses; I see a candy bar and decide it is worth pursuing, etc. But imagination and memory can initiate desire as well. Suppose I recall a particularly delicious brand of wine I drank once. I form a desire for it and decide I should set myself in

In summary, the hylomorphic state of desire (desire to quench thirst constituted by the neuronal activity) activates the faculty of practical reason in an efficient cause sort of way. Then, what gets Morgan moving, is the deliberated desire (προαίρεσις) that is the result of using practical reason. Practical reason prioritizes the desires (continue playing video games, or get up?) and combines them with various cognitive inputs to result in a more sophisticated desire (βουλευτική ὄρεξις) that acts on Morgan's limbs to get her to the fridge. Once set in motion, the change from the initial state on the couch to the final state by the fridge is guided by deliberated desire to achieve the end, in this case to quench her thirst.<sup>94</sup>

Next, it should be explained how the four causes of Morgan's getting the beer are related to the four causes of Morgan's desire for a beer.

Morgan's desire to quench her thirst is the efficient cause of the efficient cause of Morgan's eventual getting the beer, i.e., of the activation of the faculty of practical reason and of its eventual output (a deliberated desire). The initial state of desire to quench thirst (brought about whatever relevant bodily states) is what the faculty of practical reason picks up on and results in a deliberated decision. Morgan at the time of the initial state of desire has quite a bit to think about. Is getting the drink now worth it given her desire to continue playing her game? If so, what should she

---

motion to get it. Certainly some kind of perception is in play when entertaining a memory, even if it is not direct, immediate perception through the senses.

<sup>94</sup> Aristotle spends quite a bit of time working out how animals can move themselves. His theory, expounded on in *De Motu* 701b12-703a23, is very broadly that when objects of some kind of desire (sensual, intellectual, etc.) are perceived (directly through the sense, recalled in imagination, etc.), there is an accompanying heating or cooling process of the *pneuma*. This heating or cooling affects the shape of the *pneuma*, which in turn, causes the limbs to move. I am footnoting this here rather than describing it in the main text because the biology is largely incorrect and I am offering a viable interpretation of Aristotle's hylomorphic psychological theory for contemporary usage. Later, I discuss how desire can "prick" our sense of practical reason, according to Aristotle. However, there are still some key insights we can take from Aristotle. First, the faculties involved in movement still require the body, or more accurately some part of the body. Second, as with the pathos of the soul, an account of what the faculty is essentially requires the materials used to structure it. For example, suppose the subject is thought. Thought cannot be abstracted from the brain without getting errors in reasoning in how and when thought occurs (Aristotle discuss a possible exception to this: active nous. But this is beyond the scope of this dissertation). In the example of Morgan getting up, her desire for a drink would come with a heating of her *pneuma*, which, if she allows, can result in her limbs moving. At any rate, for an informative read on Aristotle's account of animal movement and desire, see: Charles (2012, pp. 75-93).

drink? Again, through the faculty of practical reason, Morgan will prioritize these desires and come to a decision, in this case settling on a beer and setting herself in motion to get it. As in the non-deliberative case, the formal and material causes concern the mechanisms by which the efficient cause plays out.

The (structured) material causes of Morgan's initial state of desire to quench thirst (constituted in neurons) are (efficiently) causally connected to the material causes of Morgan's getting a beer (the muscles, tendons, neurons, etc. that make up her arms and legs, and so on). These materials are in turn structured and sequenced such that they can bring about the next series of structured and sequenced motions that result in the end state. The deliberated desire sets the muscles/tendons/etc. in motion that result in the shape of her starting to get up at  $t_1$ , continuing to get up at  $t_2$ , finishing getting up at  $t_3$ , moving one leg in front of the other, while one hand moves back at  $t_4$ , until  $t_n$  when she is standing by the fridge in a drinking form. Care should be taken here. While this conveys the right idea about how the various parts are suitably connected, this does not all happen at the level of the material. The materials constitute the various changes that occur from beginning to end. Morgan's initial desire (constituted by neurons) is able to activate practical reason (a faculty constituted by the relevant material, including neurons). This sets in motion a chain of events that will result in Morgan's body moving. This brings us to the formal causes.

As always in Aristotle's hylomorphism, forms are constituted by suitable materials. The initial form is the desire to quench thirst, which eventually ends in the form of standing by the fridge drinking the Pabst Blue Ribbon. The sequence of events at this level would look something like this. In Morgan's initial state, she has a desire to quench her thirst and also a desire to continue playing her video game. This gives rise to/sets in motion a process of deliberation about whether or not to

pursue a drink or continue to play her video games, about what to drink, etc. This gives rise/results in a deliberated desire for the Pabst Blue Ribbon and the best means by which to achieve this (getting up and walking there). This results in the actions themselves, *actually* getting up and moving toward the fridge. At each of the stages the form of the mental state (constituted by relevant materials) has the efficient causal power to impose the next form in the sequence. Notice also that each form includes a higher-order state: desire to quench thirst, deliberating about the best way to proceed in this circumstance, settling on a course of action, and finally walking there...each of these sorts of states are constituted by the relevant materials. It remains to discuss the final causes.

The final cause of Morgan's initial desire is thirst quenching. There are a few points to make about this. First, thirst quenching, like retaliation or a completed house, is a goal. When accomplished, they are the ways in which parts of the world are changed. This is different than the desire for these states of affairs; the desire is only the *wish* for a certain state of affairs. This is important because the final cause will not be the way in which Jaworski's rationalizing cause will be mapped onto Aristotle's metaphysic. Second, this initial desire can undergo a change into a deliberated desire. Morgan, through the faculty of practical reason, contemplates the options and settles on a deliberated desire--a desire for the Pabst Blue Ribbon beer. Initially, the faculty of practical reason somehow *inherits* the initial desire of quenching thirst and results in the deliberated desire. The desired end is the drinking of the beer to better quench her thirst.<sup>95</sup>

---

<sup>95</sup> There could even be a further change of desired ends. Suppose Morgan's roommate, Ruta, thought it would be funny to drink the Pabst herself and refill it with Kool-Aid. Morgan's journey to the fridge would not result in her deliberated end being achieved. She might feel frustration or anger at Ruta. But upon further reflection, she realized that drinking Kool-Aid still satisfied in some way her initial desire of quenching thirst, even if it wasn't in her preferred way--through a nice, cold Pabst.

With this in place, there are a few considerations about the deliberative and non-deliberative causes that merits discussion. This is the subject of the next sub-section.

#### *4.1.2.3 Some Further Thoughts on Aristotle and Desire in Action*

It might seem strange that a given desire might either directly act as an efficient cause of motion or set in motion the faculty of practical reason. However, this is plausible. Consider a few of the complexities of human action. Things could have gone very differently in the initial case where Morgan immediately struck Barney in the nose. For example, Barney could have punched Morgan (driven by a desire for something). Morgan, as a trained mixed martial artist immediately starts to throw a punch. Half way up, *she stops herself* and deliberates on the best course of action. She might decide (quickly) that asking Barney why he did that would be the best initial move. She does. They talk. Turns out Barney is having a really difficult time at home and does not know how to properly express his emotions or vent his rage. Morgan invites him to her gym where he learns discipline, self-respect, and honor...and the story could go on. Under what conditions does a desire bypass practical reason and go straight to action? Under what conditions can the desire for practical reason "intervene" and stop an action? Under what conditions does a desire go straight to practical reason?

There are further questions too. Is the faculty of practical reason one capacity, or several? Are desires individuated by aims, or is there one faculty of desire? Aristotle does not go into many of these questions, and I think this is likely for the best. Desire in action is very complicated and trying to sort through all these questions would require a great deal of work from many disciplines. Cohen (2016) says something similar:

"Aristotle displays some hesitation in his discussion of desire and its relation to practical reason in the aetiology of animal action. Some have consequently concluded that his treatment can be regarded as at best inchoate or, worse, as positively befuddled. There seem to be no grounds for any such harsh assessment, however. Equally likely is that Aristotle is simply sensitive to the complexities involved in any approach to the intertwining issues in the philosophy of action...he evidently appreciates that the data and phenomena in this domain are unstable, wobbling and retreating at the approach of taxonomizing theory. The antecedents of action, he rightly concludes, involve some sort of faculty of desire; but he is reluctant to conclude that desire is the sole or sufficient faculty implicated in the explanation of purposive behavior."

I think the key insights are enough to be getting on with. Desires are constituted by suitable materials, and these materials, when suitably structured can set the animal in motion, whether immediately, or through some sort of deliberative process.

#### *4.1.2.4 Mental-to-Mental Causation*

In the first chapter, one of the exclusion problems had to do with mental-to-mental causation and the thesis of Mind-Body Supervenience. Suppose Morgan is thinking about philosophy. This thinking, in turn, causes her to feel angst. Thinking (M+) supervenes on a particular neural state (P+). The subsequent angst (M) also has a neural base (P). However, it appears that P+ is sufficient for P and so Causal Exclusion rules out M+ as a cause for M (angst).

Aristotle's view has a way out. Because M is constituted by P, and because the way something is structured has a direct role to play in a subsequent state, the hylomorphic M/P brings about the angst M+/P+. The initial M/P state acts on neurons, yes. But the structure of the M/P event, plays a big role in how the

subsequent state is structured. The upshot is that M/P causes Morgan's angst in virtue of structure and matter. And it also falls out of Aristotle's view that M and M+ would not exist at all if they were not enmattered.

With this in view, I turn now to why the causes do not compete in cases of mental causation.

#### 4.1.3 Why the Causes Do Not Compete in Cases of Mental Causation

In this section, I discuss why the four causes are not in competition in the case of mental causation. That is, I discuss why each of the four causes does not take up the work of one of the other causes. I first give a brief overview of why the causes are not in competition from using material from chapter 3. I then show why the four causes do not compete in Morgan's initial state of desire, and why they do not compete in subsequent action. It will be fairly clear why the efficient and final causes are not competition. The formal and material causes are not in competition either, but this will take longer to explain.

A brief review of what the four causes explain is useful. Each of the four causes answers the question of why (*διὰ τί*) something is as it is. The formal cause explains change; how something changes into something else. The material cause is what is changed and constitutes the form. The efficient cause explains what set the change in motion. The final cause, when present, is the aim or goal of the change. Why is there a bronze statue of Aristotle here? Because:

- (A) a suitable quantity of bronze was available for change (material cause);
- (B) the bronze constitutes the shape of Aristotle (formal cause);
- (C) the artisan (applying the craft of statuary) set the processes in motion that would result in a statue of Aristotle (efficient cause);
- (D) the artisan wanted to honor the father of hylomorphism.



The causes are not in competition because they explain different aspects of change. An efficient cause has to set the change in motion. If the artisan has nothing to work with, then no changes can occur. Additionally, without an efficient cause to change the quantity of bronze to a statue, the bronze will remain a lump of bronze. What begins the change is not the same thing as what is changed.<sup>96</sup>

The efficient cause and final cause are also not in competition, even though it appears they could be. After all, both seem to explain why a course of action was set in motion. The sculptor started the processes to honor the father ofhylomorphism. It seems like the final cause and efficient cause are explaining the same thing: the series of changes that lead to a statue, and so might be a case where the efficient and final causes overdetermine the statue.

In chapter 3, I motivated the distinction between final and efficient causes by appealing to examples where no final cause was present. For example, the quantity of bronze could be picked up by a tornado and thrown into a volcano, where it would be destroyed. Or it could corrode into a pile. But even here we see that an efficient cause (a tornado; water triggering the reaction of Bronze Disease through interaction with the chlorides in the copper) brings about a change. Setting changes in motion and reasons for doing so were clearly distinct and explained different aspects of change. The question is: are they distinct in certain cases of agential action? And I delayed the answer to this question for this chapter as it deals with psychological states.

They are still distinct and they explain different aspects of change. The artisan's goal is to honor Aristotle, which can only be fulfilled *after* the bronze has been changed into the statue. The end by itself does not bring about or set in motion any of the

---

<sup>96</sup> At least this is often the case, though not always. In human generation, for example, the efficient causes of human offspring are the mother and father, who are (presumably) human. Nevertheless, there are a huge variety of changes that occur between the union of the gametes and a fully-grown adult of the same species. The point is that an efficient cause still requires something to act on. In this case certain biological materials.

changes. But the artisan's *desire* for honoring Aristotle (constituted by neurons) can set things in motion. The goal of the statue (to honor Aristotle) does not obviate the need for the artisan's skills to start the relevant processes and to stop them when the goal has been achieved.<sup>97</sup> Further, the artisan could desire to honor Aristotle but never actually set the processes in motion for whatever reason. This does not change the fact that the artisan is in a state of desire. So the final cause helps explain why certain activities were set in motion, but it does not by itself do anything to set anything in motion--the efficient causes do that. As such the final and efficient causes are not overdetermining the end result.

The final cause in cases of the kinds of human agency involving orectic states is part of the desire, which is part of why I think it might initially seem like the efficient and final causes are in competition.<sup>98</sup> Morgan's desire for quenching thirst is very similar to the actual instance of thirst quenching.

So, the final cause and efficient cause explain different aspects of Morgan's trip to the fridge: the final cause is the end toward which she moves (quenching thirst) while the efficient cause explains the changes involved from sitting on the couch to standing by the fridge drinking beer.<sup>99</sup>

---

<sup>97</sup> Recall that the art of statuary and the artisan's knowledge are efficient causes of the statue in that they explain the required steps to change bronze into a statue. These are set in motion by the artisan's initial desire for the end.

<sup>98</sup> Aristotle also defends the idea of final causes in nature: a spider spins a web in order to catch food; a tree throws down roots in order to find water (and the roots stop growing when water is found). Aristotle does not think these cases are like Morgan or the artisan; that is, trees and spiders are not *deciding* to do these things. Here there would not be a faculty of practical reason, nor even desires necessarily. Another story would need to be told about the relevant efficient cause, which are the nutritive capacities of the spider and the tree.

<sup>99</sup> Another kind of example could help here. Imagine an artist who desires people to be healthy and not smoke. The way he brings this about might be painting images of smoke-ruined lungs, or people laying in the hospital. The end is healthy people who do not smoke. The hylomorphic state of desire is for healthy people who do not smoke. But the efficient cause (either the state of desire itself or the deliberated desire) does not set in motion a series of processes that result in health; rather the efficient cause sets in motion products that depict pain and suffering. Here is a clear case where the efficient cause and the final cause come apart. Thank you to Jamie Dow for discussion on this.

Neither is the efficient cause in competition with formal or material causes taken together. The form simply specifies what the arrangement of certain material causes constitute. By itself, this says nothing about how such an arrangement comes to be. The efficient cause is what imposes this structure and accounts for the change.

In chapter 3.3, I argued that the four causes don't compete in accounting for a non-mental entity. It was clear that the efficient, final, and material causes have distinctive jobs to do, but there was a *prima facie* worry about whether or not the material cause excludes the formal cause.

The same is true here. The efficient cause of Morgan's desire for a beer is what brings it about, the material cause is the matter it's made out of, and the final cause is what it aims at (quenching thirst). But we might wonder whether there's any distinctive work for the formal cause to do in accounting for Morgan's desire. Don't we get the form "for free" once we have the material cause?

We can give the same line of reply to this worry as we gave to the worry discussed in 3.3. The answer is that form still has distinctive role that is not squeezed out by the material cause.

Consider this passage from Jaworski (2016), who is positing a hypothetical super physicist, who:

"...possesses complete physical knowledge of the universe. [She] has complete knowledge of all the fundamental physical entities in the universe: what they are, what properties they have, what relations they stand in, and what laws govern their behaviour...[W]hen [she] describes the universe, [her] descriptions are framed solely in the vocabulary of physics—solely in terms of the characteristics of fundamental physical particles or stuffs" (p. 10).

This super physicist knows a great deal indeed! Clearly, this super physicist would know the locations, properties, and laws of all of the fundamental physical entities involved in Morgan's getting up from the couch and moving to the fridge, as well as all the laws that govern their behavior. She would know the same kind of information about the physical entities involved in Barney punching Morgan, and Morgan's retaliation.<sup>100</sup> The question is whether or not this is sufficient for explaining the formal causes. And the answer is no.

There are certain things they are blind to, in addition to those I mentioned already. Jaworski (2016) says this super physicist

“...lacks the perceptual and conceptual resources to distinguish living things from nonliving ones, or mental beings from nonmental ones. The concepts of life, perception, desire, belief, money, sex, and so on are completely beyond its ken. As a result, when it describes the universe, its descriptions are framed solely in vocabulary of physics—solely in terms of the characteristics of fundamental physical particles or stuffs. Because the super physicist does not have the concepts to distinguish living things from nonliving ones, or mental beings from nonmental ones, its descriptions make no mention of plants, animals, or people, nor do they mention any distinctive biological or psychological activities such as

---

<sup>100</sup> Following Jaworski, I am keeping the conversation here at the level of physics and what physics leaves out. It is a further question whether or not the resources of physics can take over the explanatory resources of chemistry, biology, etc. There is debate on whether or not this is even possible. See for example: (Hendry, 2006; Stapp, 2005). Take the biological domain. Imagine a super biologist who knew everything about the relevant biological properties, laws, and relations biological entities stand in. Would this be sufficient to capture Morgan's emotive state? No, because like with the super physicist the kind of emotive states would be left out. The super biologist would be able to tell us all about the neurons firing and the connections between them, but what would be left out would be neurons firing in a constituting-anger sort of way and a constituting another-kind-of-emotion sort of way. This is discussed in more detail a little later on. At any rate, we can keep the discussion focused on physics and form for the sake of ease of reading.

growth, reproduction, perception, or belief. Nor can the super physicist recognize the distinctions these things mark in the natural world” (p. 10).

The reason that life and non-life, perception, desire, belief, money, and sex etc. are beyond the reach of the super physicist is because all the super physicist would be able to explain are fundamental physical entities, their properties, laws governing their behavior, etc. If Morgan is angry one second, and happy the next, the super physicist would only be able to discuss the arrangements of fundamental physical entities. Happiness and anger are not properties of fundamental physical entities.

That is, understanding what sex, desires, or beliefs are requires something more in order to differentiate them from other things. All of these sorts of things are composed of, or constituted by, or have some kind of relation with the materials that make them up. Suppose the fundamental physical entities are just the 118 elements on the periodic table.<sup>101</sup> All of these elements have a certain number of protons, electrons, and neutrons; there are certain laws that govern how they interact. There are also certain properties they have, and so on. Without something more, what could individuate at this level an instance of a belief or desire?

The point generalizes to life and non-life. What would be the relevant difference between life and non-life at the fundamental physical level?<sup>102</sup> The fundamental physical entities of Morgan or Barney are no different than those floating around in deep space, half way around the world, or wherever. There are only the 118 elements (so far that we know about), their properties, the various relations they stand in, and the laws governing their behavior. The super physicist is blind to life itself in this way. These considerations, I take it, are what Jaworski has in view about the super physicist.

---

<sup>101</sup> The thesis that the elements on the periodic table are the fundamental physical entities in the universe is probably wrong. But what I will say next would work *mutatis mutandis* for whatever these turn out to be. Also, it might be of interest to the reader that four new elements have recently been added to the Periodic Table of the Elements. The number of known elements is now 118. See: Nala Rogers (2016).

<sup>102</sup> For the purposes of this project, I am leaving to one side the discussions about panpsychism and other such theories.

Returning to the purposes of my project, something more is required to individuate psychological states, to differentiate between instances of anger, sadness, and so on.<sup>103</sup> Suppose that when my neurons fire in this particular sequence, I feel sad. But when some of those *same neurons* are caught up in another sequenced event, I might feel happy. By accounting only for the arrangement of the various particles as they change order will not allow one to distinguish between one mental state and another. If many of the same neurons can be used in this case for anger, in that case for sadness, the difference is the dynamic structure. The key point is this: form is required to individuate psychological states and, often, is required to capture which bits of the physical world constitute them (such as with ethanol).<sup>104</sup>

Form can also explain why certain events are instances of anger across a variety of species (or even A. I.). An instance of anger occurs just when an organism or entity is desiring retaliation, which requires a certain kind of matter to be organized in some way. Perhaps an instance of human anger and animal anger differ: one has one particular neural structure, and the other has another kind of structure. However, both desire retaliation. Or imagine a robot that is able to feel angry--perhaps Data from Star Trek. His positronic brain is very different from our brain, but nevertheless, when a certain sequence and structure are manifested in his matter, he might well desire retaliation. The point is that the structured state of desire itself is what allows us to know an instance of anger when it is manifesting. This can be compared with the discussion of various kinds of works of art that display Aristotle: a painting, a sculpture, a bust. What each have in common is certain features of Aristotle that are instantiated in

---

<sup>103</sup> And to understand life from non-life, and so on.

<sup>104</sup> Of course, substance dualists, idealists, people who belief in some sort of immaterial soul will all have different non-material ways of individuating mental states. This project is concerned with a more naturalist account of individuation of such states. Aristotle's view of form and structure, I think, offers a great way of cashing in on this without have to leave the realm of the cosmos. This is not to say that the substance dualists, etc. are wrong. It is just an appeal to trying to account for as much as possible within the resources of the universe.

different mediums.<sup>105</sup> The key point here is that the distinctive contribution of form is required to explain commonalities between things made of different matter.

Admittedly, however, there is a shift between form as the structure of certain materials that constitute things like tables, houses, and so on, and form as desire for retaliation. This is the same point I was making earlier. With things like houses and tables, there are certain constraints placed on materials that clearly indicate when a house or table have arrived. This would hold true across worlds. But there appears to be no necessity whatsoever to this sequence of neuronal firings constituting this instance of anger. Again, I would appeal to the same two considerations I made previously: if we want a more naturalized account of psychological states, dynamic structures of certain parts of the body seems like a great candidate. Secondly, the sorts of materials that can constitute psychological states might be of a wider range.

To sum up, Aristotle's form and matter are not in competition in psychological states. The relevant sort of matter (biological materials) is potentially a particular psychological state. It becomes an actual psychological state when the relevant materials are structured in the relevant ways. Form explains what particular state it is.

Although we have come quite a way in showing how the causes are not in competition, not all worries are put to rest. Since we are discussing how a psychological state can cause an action (the issues of mental causation), there is another worry.

---

<sup>105</sup> By including a possible A.I. being into this discussion, I am opening the possibility that Aristotle was incorrect about psychological states being essentially enmattered, and that we should think of psychological states and their material realizers along similar lines as functionalism. Aristotle thinks that psychological states are essentially enmattered; in effect, biological materials are required for psychological states to occur, and the content of these states (desires, etc.) cannot be abstracted from their materials without errors in reasoning about when and how they occur. Whatever Data has is not really biological material. So, if Data were to get angry, this would be structured in very different kinds of materials. At this stage, it seems like we could abstract the content of anger (desire for retaliation) from the materials they occur in because they could occur apart from biological materials. I discuss this more in a later section. However, I will say here that even if a kind of functionalism should be included in hylomorphism, it does not hurt the general insight about form and matter and the constitution relation underpinning it.

In 4.1.2, it was shown how, on Aristotle's view, psychological states are efficient causes of actions (in non-deliberative cases), or the efficient cause of activating the faculty of practical reason, which is the efficient cause of action (in deliberative cases). Further, the material and formal causes are required for an efficient cause to do its work of changing something from one thing to another. And the final cause, when present, gives a reason for action.

But one might wonder whether or not these other causes, especially the formal and final, do not somehow exclude the content of the relevant mental states from the causal picture. Suppose we account for causation in terms of neurons (material cause), how they're structured (formal cause), and a given end such as quenching thirst (final cause)? Why do we need to say there is an entity--a mental state--that these are the material, formal, and final causes of? The thought here is why not simply stop at a structured material (hylomorphic compound) instead of calling it a desire to quench thirst? Alternately, why could we not say that we could explain action without reference to the end? The idea here is that if we have the structured sequence of neurons, this itself will eventually result in the final structured state, which, for short-hand, could be called the standing-by-the-fridge-in-a-drinking-beer shape. The longer-hand version would involve the structure of the relevant bodily materials. I think there are good responses to these worries.

As to the first issue--why shouldn't we just stop at the level of dynamic structure and not attribute a particular mental state (*e.g.*, desire for quenching thirst)--there is at least one response I can give. Morgan's neurons, firing in such-and-so a way cause Morgan to get a beer at the actual world. At some other world, the certain combination of blood and heat around Schmorgan's heart cause Schmorgan to get an alien beer. The two very different hylomorphic processes result in a very similar outcome. How can this



be explained? A very plausible answer is that these very different hylomorphic structures count as desires, and in particular, desires to quench thirst.

What these very different hylomorphic processes plausibly have in common, then, is the intentional content to quench thirst. This point is related to the one made earlier about why intentional action results in a certain series of movements and not others. For example, desire to quench thirst explains why Morgan got up, moved toward the fridge in a certain stepwise manner. Morgan did not, for instance, get up, fall down, and roll into the kitchen. If we compare this to Schmorgan, Schmorgan also did not get up, fall down, and roll into the alien kitchen. Or suppose Schmorgan has a very different composition altogether and instead of engaging in a process that looks like walking, she rolled to the fridge; or threw out some sort of pseudopod and moved to the fridge in this way. The point remains the same: Morgan and Schmorgan took very specific steps that resulted in a very similar outcome and the best way to explain this is that the intentional content of desire played a role in those outcomes. Because the structures that resulted in the final state are very different, perhaps the best way to explain the similar results is through desire. In this way, one should include the content of the psychological state in account of form.

The answer as to why we should retain the final cause in explaining the action, and not merely rely on the structured firings of neurons (or blood and heat around Schmorgan's heart), is also contained in the preceding explanation. The content of desire essentially involves an end state--in this case, quenching thirst. If we grant that intentional content is directed toward a desired state of affairs, then this is an indispensable part of the explanation of intentional action.

In summary, I have taken the key insights from chapter 3 about why the causes are not in competition and applied them to cases of mental causation. While there are still some concerns about why the structures involved in psychological states should

mirror those of non-psychological structures (such as houses), these can perhaps be *prima facie* answered on the grounds that structure is the nearest likely candidate for explaining the differences among psychological states. Moreover, I gave reasons why an account of structure should include intentional content to account for similarities across very different hylomorphic structures that serve as efficient causes of action.

#### *4.1.3.1 Why the Causes are not in Competition for Mental-to-Mental Causation*

What was discussed above also helps with the issues of mental-to-mental causation. Suppose Morgan's thinking about philosophy (M+/P+) causes Morgan to feel angst (M/P). You might think that the neural states are in competition with their higher-order structures for bringing about the subsequent physical state of angst. But this is not the case here.

The structure of Morgan's thinking had a direct role to play in bringing about the angst and explains why Morgan had angst instead of dancing or singing.

There is a question about why Morgan's thinking of philosophy caused her to have angst instead of joy or happiness. A lot of this will depend on Morgan's disposition toward philosophy, or the content of what she learned. If what was learned challenged her in some way, angst might make more sense. Additionally, thinking might likely be a part of the capacity of practical reason. Whatever the detailed story is, the mechanisms in place will be the same: structured hylomorphic states that bring about subsequent states in virtue of both having matter and form.

#### *4.1.4 How Form and Matter are Related in the Case of Mental States*

Form and matter also share a constitution relation. In this section, I am going to discuss this relation as it pertains to psychological states and then contrast this with two kinds of functionalist theories. This section proceeds in the following way. First, I discuss the constitution relation as it pertains to mental states, using both Cohen and

Aristotle's material. The point is the same: the relation of constitution seems implicit in Aristotle's discussion of actuality and potentiality. It emerges from the discussion that psychological states are still like any other sort of hylomorphic compound. Following this I will compare and contrast Aristotle's view with two kinds of functionalism, showing that while there are some similarities, Aristotle's view avoids the weaknesses of both.

Recall that Cohen (2016) thinks a contemporary philosopher might interpret Aristotle's account of potentiality and actuality as explaining a constitution relation. Recall also the word for potentiality--δύναμις--can also be translated as "potency" and even "power." This word does double-duty in *Metaphysics* Θ. Aristotle considers the primary meaning of the term to be "power,"<sup>106</sup> as in the source of change in some other thing. Once a power is activated, it is involved in a process of changing the other thing. Cohen (2016) notes that for Aristotle, the "exercise of such a power is a [κίνησις]--a movement or process." Thus, δύναμις in this sense is closely related to κίνησις and ἡρεμίᾳ, translated as rest. Fire is engaged in the process of burning paper, changing the paper. The power to change the paper occurs when the fire runs out of material or is extinguished.

The other sense of δύναμις is the main focus of this section. It is better translated as "potentiality." In this second sense, Cohen (2016) says it is not "a thing's power to produce a change but rather its capacity to be in a different and more completed state." For example, a quantity of bronze has the capacity to be in a different state, such as a statue, or a bullet, or whatever else bronze can become. A sapling has the capacity to develop into a more complete state, namely a full-grown tree. δύναμις in this sense, says

---

<sup>106</sup> Aristotle says the primary meaning of δύναμις is the: "...source of change in some other thing, or in the same thing *qua* other" (*Metaphysics* Θ.1, 1046a1, translation Tredennick's).

Cohen (2016) is related to actuality, (ἐνέργεια). Here are parts of this key passage of Aristotle in *Metaphysics* Θ.6 1048a25-27...1048a32-b9:

"Since we have now dealt with the kind of potency [δύναμις read: power] which is related to motion [κίνησις], let us now discuss actuality [ἐνέργεια]..."Actualty" means the presence of the thing, not in the sense which we mean by "potentially." We say that a thing is present potentially as [a statue of<sup>107</sup>] Hermes is present in the wood, or the half-line in the whole, because it can be separated from it; and as we call even a man who is not studying "a scholar" if he is capable of studying. That which is present in the opposite sense to this is present actually. What we mean can be plainly seen in the particular cases by induction; we need not seek a definition for every term, but must comprehend the analogy: that as that which is actually building is to that which is capable of building, so is that which is awake to that which is asleep; and what which is seeing to that which has the eyes shut, but has the power of sight; and that which is differentiated out of matter to the matter; and the finished article to the raw material. Let actuality be defined by one member of this antithesis, and the potential by the other. (translation Tredennick's).

Aristotle highlights here two kinds of potentiality and actuality: motion (κίνησις) to potentiality (δύναμις), and substance to some particular matter. As to the latter, in chapter 3, it was important to note that potentiality should be understood in a particular way. To say that the wood is potentially a table is not to say that the table, or form of the table, is lying dormant inside the wood, awaiting actualizing.

---

<sup>107</sup> This emendation is taken from Ross's translation of the same text.

As to the former, different processes can be said to be in potentiality in a few ways. First, someone who has the capacity of sight can potentially see when they are sleeping, and actually see when they awaken. In a second way, a person is potentially a scholar when they are the sort of being who could become, gaining mastery over a certain set of skills and knowledge. This can be contrasted with a scholar who possesses said mastery of certain skills and knowledge, but is currently engaged in fishing, say. In this case, the power is dormant, and becomes activated, and so is like the first case with the sighted person.

Aristotle suggests that in both kinds of cases, there is an analogy between them:

"that as that which is actually building is to that which is capable of building...and what which is seeing to that which has the eyes shut, but has the power of sight; and that which is differentiated out of matter to the matter; and the finished article to the raw material. Let actuality be defined by one member of this antithesis, and the potential by the other."

The builder actually building is engaged in a series of activities, which involve certain materials (including his body). Before he was actually building, the materials (including those of his body) were only potentially building. Similarly, the piece of wood is potentially a statue of Hermes (or perhaps a statue of whoever, or perhaps several whistles, or whatever the wood could be transformed into). When the wood actually has the form of Hermes, there is a statue of Hermes. The key point for my purposes is that in both kinds of examples, related by this analogy of potentiality and actuality, there is suitable material engaged in either activities, or in being actually a statue (or table, or whatever).

According to Cohen (2016), a contemporary philosopher "might make this point by refusing to identify the wood with the table, saying instead that the wood only

constitutes the table and is not identical to the table it constitutes.<sup>108</sup> That is, the wood constitutes the form of table. It is what receives the form.

If this is right, the wood is not identical to the table; it constitutes the table. As wood, it is still potentially a statue of Hermes, a quantity of whistles, or even another table. Wood can still be a variety of things. The *table is identical to* the composite of matter and form.

Psychological states are supposed to follow suit.<sup>109</sup> They have suitable materials that constitute them, and the composite of matter and form is what is identical to the psychological state. Consider Morgan's nervous system, and especially her brain. There are a wide variety of neurons and different combinations and sequences they can fire in. Plausibly, Morgan's psychological states are capacities she has (or part of a further capacity for desire which can admit of a variety of desires). In this way, Morgan's neurons, right before Barney punches her, could be potentially a variety of desires (as such, this is closer to the builder who is currently not building, but has the capacity for it). Once the punch arrives and Morgan registers the injustice of it all, a certain sequence of neurons constitutes a desire for retaliation. But many of the neurons caught up in her anger might have, under another circumstances, been caught up in other desires (perhaps a desire for drink, or some such). For this reason, we do not want to

---

<sup>108</sup> In context, Cohen takes the Hermes statue as the clearest example of the constitution relation.

<sup>109</sup> There could be a question here of whether psychological states are more like processes in the way building is, or if the materials that constitute them more similar to the wood of the statue of Hermes. Likely, they are more like the former insofar as psychological states are constituted by neuronal activity. However, the more critical point is that Aristotle treats them both in an analogous way. Certain materials constitute the relevant activities or things.

At least, this is a plausible interpretation of Aristotle. Cohen (2016) says that it is possible Aristotle might mean that the actual wood becomes the table, as opposed to constituting the table: "The as yet uncarved wood is only potentially a table, and so it might seem that once it is carved the wood is actually a table. Perhaps this is what Aristotle means, but it is possible that he does not wish to consider the wood to be a table. His idea might be that not only can a piece of raw wood in the carpenter's workshop be considered a potential table (since it can be transformed into one), but the wood composing the completed table is also, in a sense, a potential table."

I tend to think Cohen's interpretation is right; constitution is in view.

identify the state of desire for retaliation with the neurons themselves, as they could have been, in another sequence, constituting another state of desire.<sup>110</sup> Rather, it is this sequence of neurons firing that constitutes a desire for retaliation. The desire for retaliation is identical with the particular, patterned sequence of neurons (form + matter).

Morgan's desire to quench her thirst in the other example is likewise constituted by a certain sequenced pattern of neurons firing, and perhaps many of the same neurons that could have constituted anger are caught up in her desire for quenching thirst. Again, it is a good idea to not identify these desires with the neurons themselves, but with the form + matter of the desire. And the content of this desire explains what Morgan and Schmorgan have in common.

This constitution relation understanding of psychological states admits of many benefits. The key benefit is that it avoids the problems of identifying mental states with material on the one hand, and form on the other. A brief discussion of these benefits would be helped along by comparing Aristotle's view with two kinds of functionalism: realiser and role.

Janet Levin (2018) sets up an example with pain which we can use to spell out the differences between the two kinds of functionalist theories:

*"Pain* is the state that tends to be caused by bodily injury, to produce the belief that something is wrong with the body and the desire to be out of that state, to produce anxiety, and, in the absence of any stronger, conflicting desires, to cause wincing or moaning."

If the functional role of pain in human is played by stimulation of C-fibers, then humans can be in pain when undergoing stimulation of these fibers (Levin, 2018). But as Levin

---

<sup>110</sup> Indeed, blood and heat around the heart, in other combinations of heat and blood, might have constituted another kind of desire other than retaliation.

(2018) also says, there is a further question about what the property of pain amounts to: "is it the higher-level relational property of being in some state or other that plays the "pain role" in the theory, or the C-fiber stimulation that actually plays this role?"

The realiser functionalist identifies the property of pain with the C-fiber stimulation itself, the lower-level realizers of pain. Such an approach would indeed get around the problem of mental causation; C-fiber stimulation in humans is what directly produces belief about being in pain, wincing, and moaning. But the cost of this approach is undermining one of the chief motivations of functionalism itself, which is the multiple-realizability of pain. The thought is that pain could be realized in a variety of suitable systems. A squid, an alien, or an angel would have very different systems, but what functionalism was supposed to preserve is that "pain" can be realized in any suitable system. In realiser functionalism, identifying pain with C-fiber stimulation has the unwelcome result of making C-fiber pain the only kind of pain; after all, C-fiber stimulation cannot be what occurs in squids, aliens, or angels.<sup>111</sup>

Aristotle's view does not have this problem because while mental states are constituted by certain biological materials, they are not *identical* to them. This leaves room for multiple realizability because different suitable materials can constitute pain (or anger, etc.) much in the same way a house can take on a variety of shapes and structures.

Aristotle's view also avoids problems with role functionalism. A key problem with role functionalism is that it leaves itself wide open to the exclusion problem. If pain is identified with the higher-level role property, but is realized by the lower-level

---

<sup>111</sup> Some, on behalf of realiser functionalism, have argued that there could be more general physical similarities across species (see for example Churchland (2005)). However, it is not clear that this is a successful project (see Aizawa and Gillett (2009), for example). Furthermore, as Levin (2018) points out, even if Churchland, et al. are correct, "... it is unlikely that these similarities hold of *all* the creatures, including Martians and other hypothetical beings..."



neural events, then it seems that one can give an entire account of the pain event in terms of the realisers. Levin (2018) makes a similar point:

"But, some have argued (Malcolm (1968); Kim (1998; 1989)), if pain is realized in me by some neural event-type, then insofar as there are purely physical law-like generalizations linking events of that type with wincings, one can give a complete causal explanation of my wincing by citing the occurrence of that neural event (and the properties by virtue of which it figures in those laws)."

This would be true even if we denied Causal Closure of the physical domain (as I argued we should). It would appear that without a clear explanation of the contribution the higher-level role of pain, it looks like the realisers are doing all the causal work.

Aristotle avoids this problem because of his view of structure. Role functionalism includes the structure of neural firings in the lower-level realisers. But Aristotle would think that the structure is distinct from the physical realizers. That is, the patterns of firing cannot be identified with the materials themselves. A house is identical with the structured (formal cause) materials that constitute it. Similarly, anger is identical with the structured (formal cause) neural firings that constitute it. Form explains the change from a pile of wood and stone to a house in virtue of the structure of these materials; form explains the change from being not-angry to angry in virtue of the structure of neural events. And in cases of mental causation, it is in virtue of this structure that subsequent events are set in motion (walking to the fridge, instead of falling down, etc.). Thus, structure has a clear causal role to play in Aristotle's framework.

Thus, Aristotle navigates the Scylla of realiser functionalism and the Charybdis of role functionalism.

However, as Victor Caston (2006, 2009) points out, there are at least two other significant differences between Aristotle's view and mainstream functionalism:

- (1) Aristotle thinks that there is no way of securing the autonomy of psychology;
- (2) the psychological terms of Aristotle's psychological states cannot be replaced by Ramsey-Lewis style theoretical terms (see: Caston, 2009, pp. 38-39).

Caston (2009) explains:

"Functionalists rightly took psychological states to be necessarily enmattered for Aristotle. But in their effort to secure the autonomy of psychology, they identified the definitions of psychological states like anger with the definition of its formal component. In so doing, they departed from Aristotle himself...on his [Aristotle's] own view, the material component is as much a part of the essence as the formal component and so must be mentioned in the definition of the psychological state. For this reason, Aristotelian definitions are not the topic-neutral "logical descriptions" that functionalists wanted, which would allow the multiple realizability of psychological states and the autonomy of the psychological" (p. 38).

The reason that on Aristotle's view matter is as much a component as the formal component is because of his hylomorphic commitments: forms are constituted by matter and so it is the form + matter that is psychological state itself.

I would only add that Aristotle's hylomorphism can countenance multiple realizability, it is just that defining psychological states apart from their material

realizers is misguided.<sup>112</sup> A squid, Data the android, and a human can be in pain, but what constitutes pain are certain physical/material states, even if these are different. The key point is that psychological states are essentially enmattered.<sup>113</sup> Caston (2009) also says:

"...[O]n Aristotle's theory, desire is itself a hylomorphic composite. So even if his sample definition of anger does not explicitly include reference to matter in the formal component, it does so implicitly, in so far as there is a hylomorphic composite nested within it that contains matter itself" (p. 39).

The point I want to make is that what Data, the squid, and the human have in common is the content of desire, but this content must be constituted by some kind of matter.

The kind of multiple realizability in play, then, is not the sort hoped for by functionalists, which would mean the autonomy of the psychological. Such an autonomy would make psychological states akin to the "line and plane" examples Aristotle discussed. Jaworski (2011) makes a similar compare and contrast in explaining functionalism:

"Psychological discourse is not like a natural scientific theory, functionalists claim, but an abstract one. The mental states it postulates are analogous to the angles, lines, and figures postulated by Euclidean geometry. We arrive at Euclidean principles by abstraction, a process in which we focus on a narrow range of properties and then construct idealized descriptions of them" (p. 136).

---

<sup>112</sup> Some argue that multiple realizability is not as obviously true as it appears to be. See for example Shapiro (2000).

<sup>113</sup> Care is required here. Aristotle, as discussed, thinks psychological states cannot be successfully abstracted from their matter, as opposed to line and plane. However, Aristotle clearly thinks that other animals with different physiologies can have something like our psychological states. However, for Aristotle, whether line, plane, or psychological state, if one is to exist in the natural, sublunary world, they will have to be instantiated in some kind of matter. Let "essentially enmattered" refer to this idea.

Again, one of the reasons Aristotle has for thinking that psychological states are not like the line and plane is that it is specific changes of certain materials that account for what something is--the bricks and wood become a house; the neural events are so structured as to constitute anger, etc.

And so, the fact that psychological states are essentially enmattered does not mean a kind of multiple realizability is not possible.

This leads to Caston's second point about the motivations of functionalism, namely that psychological terms, in theory, can be replaced by Ramsey-Lewis style theoretical terms. Caston (2009) expands on this while comparing Aristotle's view to functionalism:

"In Aristotle's sample definition, the formal component of anger is specified as a "desire for retaliation or the like...and so explicitly includes a psychological state. Functionalists might not regard the occurrence of a psychological state in the definition as fatal, so long as it is treated as merely provisional. They themselves stress the "relational" character of psychological states and ways in which they must be specified reference to each other. It is for this reason that they approach such definitions holistically, in terms of an entire psychological theory, so that psychological terms can be eliminated from all such provisional definitions simultaneously, using the Ramsey-Lewis method for defining theoretical terms..." (p. 38).

Aristotle, rather, includes explicit psychological content, such as desires, into his hylomorphic theory of psychological states. And so, Caston (2009) says:

"...Aristotle does not appear concerned to "eliminate" psychological terms from his definitions, any more than he is to remove the presence of material terms from it" (p. 39).

As such, he is not concerned with two key motivations of functionalism. But identifying the form of psychological states with the intentional content (desire for something) of affective states is a good idea. One key concern for some (but not all) kinds of functionalism is whether or not Ramsey-Lewis style theoretical terms can account for the phenomenal content of mental states.<sup>115</sup> Beginning with the content of mental states avoids these problems. Of course, the issue I discussed in a previous section is relevant here; namely it is not obvious that the forms of mental states are essentially connected to matter, and it is not clear what constraints, if any, the kind of materials place on the hylomorphic structure. Recall that the materials themselves place constraints on the way a house can be shaped, and even help to determine certain properties of the house, such as keeping out the elements in virtue of their impenetrability, etc. But it is not clear that biological matter (or other kinds of suitable matter) places such constraints on the formal cause of psychological states. Nevertheless, what was said before can go here too. We can *prima facie* motivate this by appealing to a more naturalized account of psychological states. And biological materials in humans (and perhaps the materials in other organisms) are the mostly likely candidates. Certainly, more work needs doing on this--philosophical and empirical--but it still seems possible that certain materials can constitute psychological states and their content.<sup>116</sup>

---

<sup>115</sup> For discussion on this point, see Levin (2018). Her discussion the problems posed by inverted and absent qualia and the various ways functionalists have tried to respond. For a brief overview, see Jaworski (2011, p. 155).

<sup>116</sup> There is a *qualia* worry for Aristotle. We can know a great deal about the body and its properties through the sciences: the specific materials, the kinds of electrochemical reactions that occur, and so on. However, the phenomenal content of psychological states is first-personal and is something one has direct awareness of. There is something it is like to be on the beach, eat dessert, smell a favorite meal, etc. How can "third-personal" material such as neurons, etc. constitute first-personal content? Victor Caston (2002) writes on this issue. Here is a quote from Caston's (2002) abstract:

"Aristotle's discussion of perceiving that we perceive (On the Soul 3.2) has points of contact with two contemporary debates about consciousness: the first over whether consciousness is an intrinsic feature of mental states or a higher-order thought or perception; the second concerning the qualitative nature of experience. In both cases, Aristotle's views cut down the middle of an apparent dichotomy, in a way that does justice to each set of intuitions, while avoiding their attendant difficulties" (p. 751).

William Jaworski (2011) has an altogether different hylomorphic approach to this problem, wherein he denies that there is a kind of "inner mind" wherein mental phenomena occur. Rather, argues Jaworski (2011), mental phenomena are behavioral patterns, or structures: "According to hylomorphists these

In summary, we have seen how the formal and material causes have distinct work to do in mental causation. I have also compared and contrasted Aristotle's view with two kinds of functionalism, one of which falls prey to the exclusion problem (role functionalism) and one that loses one of the key benefits of functionalism, namely multiple realizability (realizer functionalism). Further, I contrasted Aristotle's position with functionalism more generally. There are two key motivations of functionalism that Aristotle does not share: the autonomy of the psychological, and the goal to replace psychological terms with Gordon-Ramsey theoretical terms. It is this latter point, the psychological content, that allows Aristotle to retain a version of multiple realizability based specifically on said content. But this does not fall prey to the exclusion problem because the constitution relation between matter and form as Aristotle uses it ensures there is a role for structure and matter to play in mental causation.

In the next section, I discuss where Jaworski's rationalizing and triggering causes fit into the overall Aristotelian picture.

#### 4.2 Mapping Jaworski's Rationalizing and Triggering Causes into the Aristotelian Framework

These questions remained for Jaworski at the end of chapter 2.

- (1) Why don't the rationalizing and triggering causes compete?
- (2) How are the rationalizing and triggering causes related?

Both questions turned critically on his neutrality about Causal Closure. What we want is for the rationalizing not only to render a rational status to an action, but also have a role in the action taking place.

If rationalizing and triggering causes are plugged into Aristotle's system, then both questions can be answered.

---

patterns include mental phenomena. Thought, feeling, perception, and action, say hylomorphists, are patterns of social and environmental interaction" (p. 311). See also: Jaworski (2011, pp. 314-324). I favor Aristotle's view.

#### 4.2.1 Placement of the Rationalizing Cause in Aristotle's Metaphysics

The rationalizing cause of an action is a psychological state (e.g., anger when Morgan punches Barney, or the decision resulting from use of the faculty of practical reason) where the psychological state is an entity with a final cause (anger or the desire to quench thirst).

If an action is rationalized, this typically means that it has a normative status of being rational. It could be rational in terms of being intelligible or justifiable with regard to the agent's beliefs, desires, and so on.<sup>117</sup> In the case of Morgan punching Barney, retaliation is intelligible; anyone seeing it would understand why she threw the right cross. But the action might also correspond with some of her beliefs, such as when someone attacks you unjustly, you are justified in striking back. In the case of Morgan's decision to get up, she decided that quenching her thirst now had more benefit than not. What both of these cases share in common is that the end they are aiming for *makes sense of* the subsequent action. For this reason, the desired ends lend a rational status to the action.

At this point, it is helpful to compare the Aristotelian view here with one of Jaworski's (2016) comments on rationalizing and triggering causes:

"As an empirical matter of fact, there appear to be at least two different ways of explaining human behavior: a way that appeals to reasons, and a way that appeals to physiological mechanisms...beliefs and desires cause or contribute to actions in one kind of way--they *rationalize* actions--and neural events cause or contribute to actions in a different kind of way--they *trigger* the muscular subsystems involved in actions" (p. 281).

---

<sup>117</sup> In a previous chapter, I also said that rationalizing cause would imply a certain awareness of these rational criteria of being intelligible or justifiable with regard to the agent's beliefs, desires, and so on.

Jaworski is saying there are at least two ways of explaining behavior: through appeals to reasons (rationalizing cause) and through appeal to all the relevant physiological processes involved in the moving of muscles.<sup>118</sup> Because each of these kinds of causes reveals different kinds of explanatory factors and explanatory relations, these reveal for his hylomorphist two different kinds of causation in play. Why did Morgan get up from the couch? The physicist and/or biologist might answer this question in mechanistic terms, citing the relevant muscular subsystems and/or fundamental particle movements, etc.<sup>119</sup> The psychologist might answer the question by citing Morgan's reason: she was thirsty and wanted a drink. Both explanations give a reasonable answer to the question. And both answers reveal certain causal factors in play. Morgan's desire rationalized her getting up, while Morgan's muscular subsystems triggered the getting up.

If we apply Jaworski's modes of explanation to Morgan, we get the following two answers to the question "Why did Morgan get up from the couch?":

(1) Because she was thirsty and wanted to get a drink (rationalizing cause).

(2) Because her nervous system triggered certain muscular subsystems (triggering cause).

(1) shows that her action has a reason or purpose that she holds, while (2) shows that certain muscular contractions were set in motion by key event in the nervous and muscular systems. What was not clear from chapter 2 was how to understand the relation between (1) and (2) and why (2) doesn't compete with (1).

---

<sup>118</sup> Presumably one could also cite the movements of the various fundamental particles if they wanted to as well. The thought behind triggering causes, I take it, is a way of explaining behavior using only fundamental particles and/or things composed of them (such as muscles and neurons). Jaworski's problem was including the structure as part of the triggering cause, leaving it a mysterious matter of how the rationalizing cause is related to the triggering cause.

<sup>119</sup> As we shall see later, Jaworski runs into problems because he includes structure (or form on Aristotle's view) in his account of the triggering causes.



Aristotle has provided a way of understanding in virtue of what orectic states are rationalized; the intended aim.<sup>120</sup> Desiring these ends makes sense of the subsequent action. In this next section, I will discuss how triggering causes fit into the picture.

But for now, even if the actions are rationalized, this still says nothing about how this rational status contributes to what happens. Indeed, even if there is a certain desire for an end in play, this does not by itself prevent the triggering causes taking over the causal explanation of the subsequent action. If what is doing the causal heavy lifting is simply the neurons firing and triggering the relevant muscular subsystems, it is difficult to see what causal role, if any, a rationalizing cause can achieve. And, without denying Causal Closure, this might well be an insurmountable task.

But, as discussed in chapter 2, there is good reason to deny Causal Closure. This, by itself, does not explain the causal role of a rationalizing cause. And so here, I would like to present a view based on Aristotle's.

An orectic psychological state is a desire for an end constituted by suitable materials (neurons in this case). The form of the hylomorphic process is the desire for the particular end (revenge, thirst quenching, etc.) while the matter that constitutes it is the structured firing of neurons. The sort of desire in play turns critically on how it is constituted--i.e., the particular structure of neurons firing. What role does desire so constituted play in what happens next? The answer is that it shapes the next stage of action: Morgan didn't tap dance when Barney punched her, she threw a right cross. She didn't start doing push-ups, she threw a right cross. The desire for retaliation in virtue of its structured sequence helps to determine each subsequent stage of events.

Similarly for Morgan's decision to get up from the couch. The reasoned desire to get up is constituted by a certain sequence of neurons firing. This in turn determines

---

<sup>120</sup> Indeed, this is a big part of why a hylomorphist can hold Anti-reductionism (mental properties are distinct from and irreducible to physical properties): a mental state has a final cause.

why she got up, moved across the living room, and opened the fridge, etc., and not fell over, or walked upstairs, or started swinging a shoe around in circles by the shoelace. The content of desire, constituted in neuronal firings, thus plays a critical role in subsequent action.

The rationalizing cause, when mapped into Aristotle's view, provides not only a mechanism by which an orectic state is rationalized (the desired end), but also the hylomorphic nature of orectic states provides a way for the content of desire to play a direct role in bringing about subsequent action.

The Aristotelian view on offer does not drive against Jaworski's preferred picture, except insofar as it requires a breaking of the neutrality on Causal Closure. Why does Aristotle's view require the denial of Causal Closure? After all, the desire itself is constituted by physical stuff. This goes back to the metaphysics in chapter 3. While forms are constituted by materials, they are not identical to them.

And to the extent that form is not completely identical with the physical materials that constitute it, is the extent to which it is not part of the physical domain. Hylomorphism still requires the rejection of Causal Closure even though physical matter is what constitutes psychological states. The reason for this is because it is still the case that Morgan's decision and aims are why certain matter is set in motion. And decisions and aims are not properties of fundamental physical matter. A biological example can help too. A cell is constituted of a colossal number of atoms. These atoms constitute a myriad of functions and things. Mitochondria help direct processes that lead to usable chemical energy; tRNA, with the help of RNA polymerase, translates codons from mRNA into amino acids used in making proteins. It does not seem like making proteins or usable chemical energy are properties that fundamental physical entities would have unless the structures that are constituted by them conferred certain powers

in virtue of said structures.<sup>121</sup> Remove or destroy the structures and you remove or destroy the functions; but the materials would remain. If this is right, then in Kim's terms, biological laws "inject causal influence" into the physical domain.<sup>122</sup>

Further, even though it is not a full member of the physical domain, it is being afforded a causal role in events in the physical domain (punching, walking, etc.), and specifically in how subsequent actions are structured. And, as we saw, the forms themselves play a role in determining which sorts of actions come next.

But I think, too, this sort of picture gives Jaworski a better explanation for understanding precisely how rationalizing causes contribute to action (at the cost of Causal Closure). For example, Jaworski sets up the following possible counter-argument an opponent might level against his claim that rationalizing and triggering causes do not overdetermine their effects. This opponent, says Jaworski (2011), argues:

"Suppose, for instance, that your act of reaching for a bottle of water is rationalized by a desire and is also triggered by an event in your nervous system. In that case, it looks like your action is overdetermined--not by two rationalizing causes or by two triggers, but by the one rationalising cause together with the one trigger" (p. 347).

The hylomorphist, the opponent concludes, must either reject the rationalizing cause or the triggering cause. But, continues Jaworski (2011), the hylomorphic solution would then "[i]n the former case...collapse into a form of epiphenomenalism, and in the latter

---

<sup>121</sup> Can the higher-order biological laws and operations be explained through the resources of physics? Are these higher-order properties explainable via the properties and relations of the fundamental physical properties? Of course, these are open empirical and philosophical questions at this point. I am arguing for the plausibility of not identifying structures with physical entities.

<sup>122</sup> It is important to note that the issue here is not about necessitation: that is, one prior state necessitating a subsequent state. Even if causation laws in physics turn out to be probabilistic, the issue is whether or not the explanatory resources of physics are sufficient for other domains. Here again is Kim's (2011) quote from chapter 1: "To deny this [Causal Closure] would imply that there might be non-physical causal agents, outside the physical domain, injecting causal influences into it, and that there might be physical events whose explanations must invoke these non-physical causal agents and forces. *If this were the case, theoretical physics would be in principle incompleteable.* Note that Causal Closure understood this way does not exclude non-physical causes of physical effects; the closure principle only says that if there are to be such causes, there must also be sufficient physical causes" (p. 38) (italics mine).

case...end up endorsing *the violation of the physical laws* the way some emergentist theories do" (p. 347) (italics mine).

I will turn to the italicized portion briefly, but Jaworski (2011) argues that the hyломorphist can insist that actions "...have rational structures; triggers, on the other hand, are causes at a lower level of organization; they need not fit into higher patterns of reasoning" (p. 347). The ultimate picture that Jaworski (2011) wants to insist on is that "a complete explanation of human action involves more than just the triggers of the bodily movement; it involves rational, social, and environmental factors as well..." (p. 348). Jaworski's second argument against his opponent is that rationalizing and triggering contribute to actions in different ways; a consequence of his causal pluralism. As such, they are not competing for sufficient cause.

But I don't think Jaworski is getting the concern in the right place. If we tweak the opponent's argument away from overdetermination, to an argument that focuses on competition, we can see it:

Suppose, for instance, that your act of reaching for a bottle of water is rationalized by a desire and is also triggered by a structured events in your nervous system. How does a rationalizing directly contribute to picking up the water bottle? It seems like the story could be told using the resources of physics and bodily mechanisms.

And the point here generalizes to his social and environmental structures as well. It seems like even social actions (like voting) could be explained in terms of physics. So insisting that rationalizing causes rationalize actions is not yet to say anything about how to assuage the exclusion problem.

Aristotle, as I've argued, has a way of explaining how rationalizing causes contribute to bringing about actions. Earlier, I italicized in Jaworski's quote the idea that certain emergentist theories violate the laws of physics. While I do not think Aristotle's

view violates any physical laws, it certainly injects non-physical causal influence into the physical domain (to use Kim's turn of phrase). Nevertheless, by accounting for how rationalizing causes are related both to their physical base and to their effects, it provides a way for Jaworski to get the sort of picture he wants. The pitcher's throwing of the baseball is not just rationalized, but the content of the desire shapes in a very real way the sequence of events that leads to the ball exiting her hand. Without denying Causal Closure, it is hard to see how the structured triggering causes in play are *not* sufficient for the throw (or indeed the punch, or the walking to the fridge). Indeed, the exclusion problem remains in force.

At this point, I turn to where the triggering cause maps onto Aristotle's framework.

#### 4.2.2 Placement of the Triggering Cause in Aristotle's Framework

Jaworski's triggering cause of an action would not just be the lower-order physical properties (neurons, etc.) in play, but rather the triggering cause would be the mental state itself as an entity with matter and form (patterns of neural firings). Again, this is something that goes against Jaworski's (2016) expressed view of the triggering cause:

"...[N]eural events cause or contribute to actions in a different kind of way--they trigger the muscular subsystems involved in actions" (p. 281).

As seen previously, for Jaworski, these neural firings activate the relevant muscular sub-systems, etc. in the coordinated ways that result in a baseball throw, or Morgan walking to the fridge, or Morgan punching Barney.

But, on Aristotle's view, such a view of the triggering cause *includes* the formal cause in the material cause. It is specifically the coordination and sequence of these neural firings that is the desire to throw the baseball, or the desire to quench thirst, or the desire for retaliation. As such, Aristotle would identify the triggering cause with the

mental state itself, which is constituted by a certain sequence of neural firings. The sequenced structure plays a critical role in what happens next.

This helps bypass a key tension in Jaworski's account between the rationalizing and triggering causes. Recall that in the previous section it was a problem for Jaworski to explain just how a rationalizing cause accounts for subsequent action.<sup>123</sup> Rationalizing the action is not enough to explain a causal role. Here we can see plainly that a rationalizing cause would be mysterious on Jaworski's account; namely, he includes the coordinated structure of neurons and muscular subsystem in his account of the triggering cause. However, on Aristotle's view the sequence and coordination of these neural firings is in the realm of the formal cause. And it is the particular structure that plays a critical role in what happens next.

Let this suffice for the placement of the triggering cause: it is the mental state as an entity with matter and form. In the next section, I discuss how the triggering and rationalizing causes are related to one another.

#### 4.2.3 How the Rationalizing and Triggering Causes are Related, and Why They Are Not in Competition

In the previous two sections, the discussion was about where the rationalizing and triggering causes fit into Aristotle's picture. The rationalizing cause of an action is a mental state with a final cause, while the triggering cause is the mental state as entity with matter and form.

The relation between the two is *identity*. The mental state is that which both triggers and rationalizes the action. The mental state is constituted by the sequence of the neurons firing and both the form of the psychological state (desire for revenge, quenching thirst, etc. with their associated goals) and the material play critical roles in

---

<sup>123</sup> As well as how a rationalizing cause is related to the neural/muscular base.

the subsequent changes. The mental state is the efficient cause, therefore, of the actions that come next.

Recall from chapter 3 that the efficient cause is in play at each stage of a coordinated effort, such as building the house. The efficient cause is at work when the builder now saws the wood, now nails the planks together, now puts the roof on, etc. The forms of each state play a critical role in what happens next: the shape of the body in hammering in nails plays a critical role in why the nails are sunk into the wood. In a similar way, insofar as Morgan is in the rationalizing state of pursuing her goal, each of her subsequent movements are directed toward the end. The structure of her neural states plays a critical role in each subsequent stage as she moves to the fridge.

While moving from the couch to the fridge, Morgan goes through a variety of shapes; many, many forms come to be and pass away, each of which comes to be and passes away via the rationalizing and triggering causes at work. The same kind of story also holds at the neurobiological level. Morgan is considering whether or not it is worth pursuing the beer; this faculty is itself constituted by biological materials. Once Morgan decides to pursue the beer through this faculty, she sets herself in motion. The change that is explained is that of not reasoning to reasoning.<sup>124</sup> The considered desire that is the result of the faculty of practical reason is the efficient cause of setting in motion that sequence of events that leads to Morgan going to the fridge. This considered desire is constituted by sequenced neural firing both the structure and material of which determine the subsequent stages of movement.

This takes away any coincidence with Morgan's desire and the successful achieved result because the structure explains why she moved in *this* way instead of, say, rolling into the dining room, or tap dancing to the bathroom. Morgan, when getting

---

<sup>124</sup> This is the same kind of case Aristotle used earlier when Aristotle was describing the potentially awake person becoming actually awake when her eyes open.

up from the couch to go to the fridge, moves fluidly and in a coordinated fashion. The rationalizing and triggering cause--the mental state--are coordinated in such a way that they align with her goal of reaching the fridge. So, the triggering causes are often sequenced in accordance with the goal the subject is pursuing; thus, the triggering causes in these instances are also structured, and this means there are formal causes in play. And so the efficient cause of the action, the mental state, contributes to what happens next in terms of both its matter and form.

In cases where the faculty of practical reason is in play, the causal story looks like this, borrowing some of Jaworski's terminology. There are two activity-making structures in play: Morgan's reasoning about whether or not to get up, and Morgan's actual getting up. Morgan's reasoning will be constituted by a certain kind of neural activity at a particular time. (What prompted, or set in motion, this reasoning was her state of desire, which was prompted in turn by her physiological condition.) Morgan, via her deliberated desire, sets herself in motion, getting up and heading to the fridge. Her actions are rationalized in virtue of having a goal or end.

A similar story holds for Morgan punching Barney as well, but the faculty of practical reason is bypassed. The desire for retaliation, constituted by neurons, sets in motion her fist, which lands on Barney's nose. The rationalizing and triggering elements of the same mental state explain why she threw the punch instead of tap-dancing, or singing a dirge, or sipping her milk.

In summary, the relation between rationalizing and triggering causes is identity. And both the matter and form have a critical role to play in action. The structure constituted by the material cause plays a critical role in what specific actions happen next. This bridges the gap between Jaworski's initial account of rationalizing and triggering causes and shows how they are essential to action.



The identity of the triggering and rationalizing causes gives us an elegant answer to the worry about competition.<sup>125</sup> They cannot compete with each other, simply because they are one and the same entity. Where Jaworski went wrong was including the structures of materials into the triggering cause (neural and muscular activity) which leaves no room for the rationalizing cause (the mental state) to play a role in subsequent events. The claim that mental properties are distinct from, and irreducible to, physical properties is compatible with the view we've arrived at through Aristotle: (a) physical matter constitutes a form that is distinct from it, and (b) this hylomorphic composite is a mental state that both triggers and rationalizes actions and other mental states. Since the mental state involves the form, which is not reducible to physical entities, the mental state isn't reducible to physical entities either.

---

<sup>125</sup> I would like to thank Heather Logue for discussion on this paragraph.

## Chapter 5: Conclusion

In what follows, I am going to summarize the discussion from each chapter. The conclusion is that Aristotle's version of hylo-morphism solves the exclusion problems for mental causation in an elegant and systematic manner.

Chapter 1 discussed the exclusion problems (mental-to-physical causation and mental-to-mental causation) for mental causation according to Kim. The mental-to-physical causation problem is the one by which a mental cause seems excluded from bringing about a subsequent physical state. For example, Morgan, desiring to quench her thirst, gets up from the couch and starts walking. Anti-reductionism says that Morgan's state of desire is not reducible to the neurons that constitute it. The exclusion problem was that, given Causal Closure of the physical domain, and Causal Exclusion, it would seem the sufficient explanation of Morgan getting up was her prior neural state. Causal Closure ensures that getting up has a sufficient physical cause, namely the neurons that set everything in motion. Causal Exclusion then ensures that there could be no other sufficient causes in play. As such, the mental content of Morgan's desire is *excluded* from contributing to her getting up since it is a non-physical influence and the physical domain is causally closed.

The mental-to-mental causation problem is related to the previous. It applies to any view of Anti-reductionism that also holds a supervenience thesis. It would seem that Morgan's thinking about the philosophical lesson supervenes on a particular physical state. Her subsequent angst also supervenes on a physical state. However, Causal Closure kicks in and says that the sufficient cause of her angst is the prior physical state that thinking about the philosophy of mind lesson supervenes on. But this prior physical state causes the subsequent physical state that her angst supervenes on. There is no mental-to-mental causation here. The cumulative force of the first exclusion

problem and this one is that not only do mental causes not cause subsequent physical states, but that they themselves are epiphenomenal to physical states. Further analysis reveals that either Causal Closure has to go, or Anti-reductionism does.

Anti-reductionism is a strong thesis in its own right. There are good reasons to think that mental states are related to their effects in some way. For one thing, it sure seems like an amazing coincidence that whenever certain desires (and other mental states) are followed by actions that result in satisfying the aims of the agent. For example, Morgan's desire to quench her thirst is coincidentally united with the physical events that bring about her getting up.

On the other hand, we do not want to say her getting up is overdetermined either: we do not want to say that Morgan's getting up has a sufficient mental cause and a sufficient physical cause. Systematic overdetermination is a bad result. Recall the baseball pitcher. We do not want to say that had the physical events not occurred, the baseball pitch would still have happened.

What we want is an account of how the mental state could contribute in some way to the baseball pitch (and Morgan getting up). But whatever contributions we would want to say that mental states contribute will always be excluded from the causal picture as long as Causal Closure holds.

If we want mental states to contribute to physical effects, and if we want to hold that mental states are not reducible to physical states, then we have to deny Causal Closure.

In chapter 2 I gave reasons based on Steward and Cartwright's work why rejecting Causal Closure of the physical domain is plausible. There are good reasons to do so. For one thing, Causal Closure seems heavily influenced by the historical hope of determinism and causation-as-production. But there are good reasons to think that, even in physics, determinism does not always hold. To use Steward's example again, the

trajectory of a tennis ball could be interfered with by a number of physical factors. Moreover, causation-as-production sorts of causes have competitors, such as mental causes. Additionally, there are debates in the special sciences, such as chemistry and biology to the effect that chemistry and biology are not reducible to physics. However, both biological and chemical factors determine to some degree what happens in the physical domain. If they are not, in fact, reducible to physics, then they "inject" causal influence into the physical domain. The cumulative force of these considerations is that the strength of the thesis of Causal Closure is in serious question. For these reasons, I have proceeded as if Causal Closure does not hold.

However, denying Causal Closure is only part of the battle. While it opens the door for different, non-physical and non-causation-as-production causes to contribute to effects, it is unclear how they are related both to each other and their effects. To this end, I introduced Jaworski's causal pluralism and his account of hylomorphic mental causation which features rationalizing and triggering causes. However, even though Jaworski makes a lot of progress with his version of hylomorphism, Jaworski attempts to remain neutral on whether or not Causal Closure holds. As a result, he is wide open to at least the mental-to-physical exclusion problem. However, his insights into a rationalizing cause and a triggering cause are promising; they just need a suitable causal metaphysic.

This is where Aristotle helps immensely. In chapter 3 showed that Aristotle argues plausibly for four distinct sorts of causes and how they contribute to causation. Matter constitutes forms. But matter and form are distinct. For example, a quantity of bronze could become bullets, a statue, bowls, or other things. One worry for this view is that form is reducible to matter. But this is not the case because form explains why something is what it is, and why certain things have certain distinctive properties they would not have unless it was so structured. For example, dimethyl ether and ethanol

have exactly the same material causes, but very different properties entirely in virtue of their structure. What something is is identical with its material and formal causes. What brings about these changes? The efficient cause explains this. The efficient cause is what acts on matter to change it into something else. This is the cause that is most akin to our notion of causation-as-production. It directly acts to bring about changes. The final cause, when present, explains why a certain change was set in motion. In cases of agential causation, the final cause explains the end for which action is undertaken.

In chapter 4, I discussed both how Aristotle applies his causal pluralism to mental states, and to how these mental states feature in mental causation. Aristotle treats mental causation in the same way. Aristotle's causal pluralism and hylomorphism give us the resources to explain (1) how the psychological states are related to the nervous system, and (2) the role psychological states play in bringing about their effects. I then map Jaworski's rationalizing and triggering causes onto Aristotle's metaphysic.

How does Aristotle's metaphysic address the two exclusion problems of mental-to-physical causation and mental-to-mental causation? The answers to both are similar. The key insight is that a mental property is constituted by a physical property. The mental property is not identical to the physical property. Rather the mental and physical properties are related as form and matter via a constitution relation. In Jaworski's terms, the rationalizing cause is constituted by the triggering causes. The psychological state is identical to the form/matter compound.

How could a mental cause contribute to an effect? If a rationalizing cause and a triggering cause are the hylomorphic compound of her decision to get up, then she stands up in virtue of having a reason constituted by her neural states. The structure helps determine which precise action the triggering causes will start. It is in virtue of the structure of her mental state that explains why she got up and aimed at her goal, rather than falling down, doing nothing, or whistling.

A similar treatment is given for the problem of mental-to-mental causation.

Suppose Morgan is thinking about philosophy and this in turn causes her angst.

Thinking about philosophy rationalizes her angst. That is, before she learned about the problem of mental causation, she was secure in the belief that there were minds and/or mental properties. After the lesson, she began to experience angst: butterflies in her stomach, wringing her hands, etc. The *reason* for this behavior was her challenged belief. And her thinking about it in the first place started the process that led to angst. Thinking is constituted by neurons, and so we can call this state R/T (for rationalizing and triggering cause). The structure of this state plays a direct role in shaping the hylomorphic structure of the subsequent state of angst and why she experiences angst and not jubilation or ennui.

The key to the solution is that structure plays an important role in causation and is not reducible to matter that constitutes structure. And so, Morgan closes her copies of Aristotle's works, having moved from melancholia to eudaimonia.

## References

- Aizawa, K., & Gillett, C. (2009). The Multiple Realization of Psychological and Other Properties in the Sciences. *Mind and Language*, 24(2), 181-208.
- Annas, J. (1982). Aristotle on Inefficient Causes. *Philosophical Quarterly*, 32(192), 311-326.
- Aristotle. (1933). *Metaphysics Books I-IX* (H. Tredennick, Trans. Vol. Aristotle XVII). Cambridge, Massachusetts; London, England: Harvard University Press.
- Aristotle. (1955). *On Sophistical Refutations; On Coming-To-Be and Passing-Away; On the Cosmos* (E. S. Forster & D. J. Furley, Trans.). Cambridge, Massachusetts; London, England: Harvard University Press.
- Aristotle. (1957). *The Physics Books I-IV* (P. H. Wicksteed & F. M. Cornford, Trans. Vol. Aristotle IV). Cambridge, Massachusetts; London, England: Harvard University Press.
- Aristotle. (1984). *Metaphysics*. In J. Barnes (Ed.), *The Complete Works of Aristotle* (Vol. II). Princeton, New Jersey: Princeton University Press.
- Aristotle, & Bostock, D. (2008). *Physics: A New Translation by Robin Waterfield* (R. Waterfield, Trans.). Oxford: Oxford University Press.
- Aristotle, & Hett, W. S. (1995). *On the Soul Parva Naturalia On Breath with an English translation by W. S. Hett* (W. S. Hett, Trans. Vol. VIII). Cambridge, Massachusetts; London, England: Harvard University Press.
- Baker, L. R. (2011). Non-reductive Materialism. In B. P. McLaughlin, A. Beckermann, & S. Walter (Eds.), *The Oxford Handbook of Philosophy of Mind* (pp. 109-127). Oxford: Clarendon Press.
- Cartwright, N. (1999). *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

- Cartwright, N. (2002). *How the Laws of Physics Lie*. Oxford and New York: Clarendon Press.
- Caston, V. (2002). Aristotle on Consciousness. *Mind*, 111(444), 751-815.
- Caston, V. (2006). Aristotle's Psychology. In M. L. Gill & P. Pellegrin (Eds.), *A Companion to Ancient Philosophy* (pp. 316-346). Oxford: John Wiley and Sons Ltd.
- Caston, V. (2009). Commentary on Charles. *Proceedings of the Boston Area Colloquium in Ancient Philosophy*, XXIV, 2008, 30-47.
- Churchland, P. (2005). Functionalism at Forty: A Critical Retrospective. *Journal of Philosophy*, 102, 33-50.
- Cohen, S. M. (2016). Aristotle's Metaphysics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 ed.).
- Contessa, G. (2014). One's a Crowd: Mereological Nihilism without Ordinary-Object Eliminativism. *Analytic Philosophy*, 55(2), 199-221.
- Davidson, D. (2002). Mental Events. In *Essays on Actions and Events* (Second Edition ed., pp. 207-227). Oxford: Clarendon Press.
- Falcon, A. (2015). Aristotle on Causality. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.
- Freeland, C. A. (1995). Accidental Causes and Real Explanations. In L. Judson (Ed.), *Aristotle's Physics: A Collection of Essays*. Oxford: Clarendon Press.
- Gardiner, H. N. (1919). The Psychology of the Affections in Plato and Aristotle. *The Philosophical Review*, 28(1), 1-26.
- Grupp, J. (2006). Mereological Nihilism: Quantum Atomism and the Impossibility of Material Constitution. *Axiomathes*, 16(3), 245-386.
- Heil, J. (2004). *Philosophy of Mind: A Contemporary Introduction* (Second ed.). New York and London: Routledge Taylor and Francis Group.



- Hendry, R. F. (2006). Is there downward causation in chemistry? In D. Baird, E. Scerri, & L. C. McIntyre (Eds.), *Philosophy of Chemistry: Synthesis of a New Discipline* (pp. 173-189). New York: Springer-Verlag New York Inc.
- Irwin, T. (1988). *Aristotle's First Principles*. Oxford: Clarendon Press.
- Jaworski, W. (2009). The Logic of How-Questions. *Synthese*, 166(1 (Jan., 2009)), 133-155.
- Jaworski, W. (2011). *Philosophy of Mind: A Comprehensive Introduction*. Chichester, United Kingdom: Wiley-Blackwell.
- Jaworski, W. (2012). Powers, Structures, and Minds. In R. Groff & J. Greco (Eds.), *Powers and Capacities in Philosophy: The New Aristotelianism* (pp. 145-171). New York: Routledge.
- Jaworski, W. (2016). *Structure and the Metaphysics of Mind: How Hylomorphism Solves the Mind-Body Problem*. New York: Oxford University Press.
- Jr., F. D. M. (1999). Aristotle's Philosophy of Soul. *The Review of Metaphysics*, 53(2), 309-337.
- Kalimtzis, K. (2012). *Taming Anger: The Hellenic Approach to the Limitations of Reason*. London: Bloomsbury Academic.
- Kim, J. (1992). "Downward Causation" and Emergence. In A. Beckermann, H. Flohr, & J. Kim (Eds.), *Emergence or Reduction?: Essays on the Prospects of Nonreductive Physicalism*. Berlin and New York: Walter de Gruyter.
- Kim, J. (1998). *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, Mass.: MIT Press.
- Kim, J. (2011). Mental Causation. In B. P. McLaughlin, A. Beckermann, & S. Walter (Eds.), *The Oxford Handbook of Philosophy of Mind* (pp. 29-52). Oxford: Oxford University Press.

- Kim, J. (Ed.) (1989). *Mechanism, Purpose, and Explanatory Exclusion*. Cambridge: Cambridge University Press.
- Kroedel, T. (2015). Dualist Mental Causation and the Exclusion Problem. *Nous*, 49(2), 357-375. doi:10.1111/nous.12028
- Levin, J. (2018). Functionalism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition ed.).
- Lewis, D. (1986). Causal Explanation. In *Philosophical Papers* (Vol. II, pp. 214-240). New York; Oxford: Oxford University Press.
- Lewis, D. K. (1983). Mad Pain and Martian Pain. In *Philosophical Papers* (Vol. I, pp. 122-132). New York; Oxford: Oxford University Press.
- Malcolm, N. (1968). The Conceivability of Mechanism. *Philosophical Review*, 77, 45-72.
- Marmodoro, A. (2014). *Aristotle on Perceiving Objects*. New York: Oxford University Press.
- Moravesik, J. (1974). Aristotle on Adequate Explanation. *Synthese*, 28.
- Nussbaum, M. (1978). *Aristotle's De Motu Animalium*. Princeton: Princeton University Press.
- Papineau, D. (2011). The Causal Closure of the Physical and Naturalism. In B. P. McLaughlin, A. Beckermann, & S. Walter (Eds.), *The Oxford Handbook of Philosophy of Mind* (pp. 53-65). Oxford: Oxford University Press.
- Penfield, W. (1975). *The Mystery of the Mind: A Critical Study of Consciousness and the Human Brain*. Princeton, NJ: Princeton University Press.
- Rogers, N. (2016). Four New Elements Complete the Seventh Row of the Periodic Table. *Science*.
- Shapiro, L. (2000). Multiple Realizations. *Philosophy of Science*, 97(12), 635-654.
- Shields, C. (2007). *Aristotle*. London and New York: Routledge.

Shields, C. (2014). *Aristotle* (Second ed.). London and New York: Routledge.

Shields, C. (2016). Aristotle's Psychology. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2015 ed.).

Stapp, H. (2005). Quantum Interactive Dualism: An Alternative to Materialism. *Journal of Consciousness Studies*, 12(11), 43-58.

Steward, H. (2014). *A Metaphysics for Freedom*. Oxford: Oxford University Press.

Tumay, H. (2016). Reconsidering learning difficulties and misconceptions in chemistry: Emergence in chemistry and its implications for chemical education. *Chemistry Education Research and Practice*, 17(2), 229-245. doi:10.1039/C6RP00008H

van Fraassen, B. (1980a). A Re-Examination of Aristotle's Philosophy of Science. *Dialogue*, 19, 20-45.

van Fraassen, B. (1980b). *The Scientific Image*. Oxford: Clarendon Press.