

ABSTRACT

Title of Dissertation: ON THE ACTIVITIES AND PARTS OF THE
MECHANISMS OF LIFE

Kalewold Hailu Kalewold, Doctor of
Philosophy, 2021

Dissertation directed by: Professor Lindley Darden, Philosophy

This dissertation advances the new mechanistic philosophy of science by developing novel accounts of activities and good parts. In the first chapter, I develop a Hybrid Account of activities that integrates production and difference-making approaches to causation, enabling the identification and individuation of causally productive activities. In the second chapter, my account of good parthood grounds being a good part in the role parts play in mechanisms as activity-enablers as well as their inclusion in what I call the explanatory mosaic of science. This account is robust enough to characterize parts of mechanisms throughout the life sciences. In the third chapter, I apply the account I develop to the case of the use of race in epidemiology and biomedicine. I show how the mechanism discovery approach, and the accounts I develop in earlier chapters, offer a normatively and explanatorily attractive methodology to researching, diagnosing, and treating complex trait disorders. The dissertation applies these accounts to case studies from the life sciences to show how they solve outstanding problems in philosophy and biology.

ON THE ACTIVITIES AND PARTS OF THE MECHANISMS OF LIFE

by

Kalewold Hailu Kalewold

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Advisory Committee:

Professor Lindley Darden, Chair

Professor Harjit Bhogal

Professor Peter Carruthers

Professor John Moulton

Professor Eric Sidel

© Copyright by
Kalewold Hailu Kalewold
2021

Acknowledgements

I would first like to thank my advisor Lindley Darden. I could not have completed this dissertation without her patient guidance, incisive comments, and unfailing support. Her deep knowledge of the history and philosophy of science and biology were invaluable in the development of my research.

I would also like to thank the other members of my dissertation committee— Harjit Bhogal, Peter Carruthers, Eric Sidel, and John Moulton— who provided generous comments and illuminating discussions of my work throughout my doctoral career.

If I have any capabilities as a philosopher, it is because of the philosophical community at the University of Maryland, College Park. I thank all my teachers and fellow graduate students at Maryland, from whom I have learned much. My thanks in particular to Aiden Woodcock, Julius Schonherr, Aida Roige, and Chris Masciari. My thanks also to Sonya Acosta, Oliver Jury, Jessica Mardo, Pat Cabral, and Eric Gilligan, who were my housemates for all or part of the seclusion imposed by the COVID-19 pandemic. Their friendship, companionship, and support made that difficult time bearable. Special thanks to Aylon Manor, who shared an office with me for a year and with whom I had long philosophical discussions; to Shen Pan, with whom I lived for the first three years of my program; and especially to Javiera Perez-Gomez, who was my first and closest friend at Maryland.

I would also like to thank the DC History and Philosophy of Biology (DCHPB) group— which has been an immensely rewarding intellectual community to be a part of— including Makmiller Pedroso, Nancy Hall, Dan Hicks, and other former and current members and attendees.

In the end I want to thank those responsible for the beginning, to my parents and my family, who always let me argue my point, and supported me all the way.

Parts of the third chapter of this dissertation have previously been published in Kalewold (2020) “Race and medicine in light of the new mechanistic philosophy of science” *Biology & Philosophy* 35(4): 1-22.

Contents

Abstract.....	1
Acknowledgements.....	ii
List of Tables	vi
List of Figures	vii
Introduction	2
1	8
The Hybrid Account of Activities.....	8
1.1 Introduction	9
1.2 The Activities View	13
1.3 Activities and a Lack of Explanatory Constraints	20
1.4 Difference-Making and Causal Production	24
1.5.1 A Hybrid Account of Activities as Productive Difference-Makers.....	28
1.5.2 HAA as a Superior Approach to Causal Specificity	40
1.5.3 HAA and the Right Balance Between Generality and Specificity	44
1.6 Conclusion.....	48
2	50
Good Parts and the Explanatory Mosaic.....	50
2.1 Introduction	51
2.2 The Carving Problem	52
2.2.1 Parsimony	55
2.2.2 Parts as Activity-Enabling.....	56
2.2.3 Parts Delineated by Boundaries.....	57
2.2.4 Technological Identification of Parts	58
2.3 A Novel Account of Good Parthood.....	59
2.3.1 The Explanatory Mosaic.....	64
2.4 Lineages as Sources of Explanatory Relations in Biology.....	65
2.4.1 Lineage Explanations and Good Parts.....	67
2.4.2 Connectedness and Gerrymandered Parts.....	69
2.5 Conclusion.....	71
3	72
Mechanisms, Race, and Medicine	72
3.1 Introduction	73

3.2 The Biomedical Race Debate	77
3.2.1 The Realists' Race as a Gerrymandered Part	79
3.2.2 Race and Productive Difference-making.....	83
3.3 Race and Biomedicine: Two Approaches.....	85
3.3.1 The Granularity Problem.....	89
3.3.2 The Reification Problem.....	92
3.4 Discovering Mechanisms	95
3.4.1 Assessing the Evidence	98
3.4.2 Building the Mechanism Schema.....	100
3.5 Conclusion.....	102
Conclusion.....	105
References	110

List of Tables

1.1 Kaiser's (2018) list of the main features of activities.....	17
1.2 Glennan's list of the necessary and sufficient conditions for being an activity.....	18

List of Figures

1.1 Phosphorylation of inactive glycogen phosphorylase.....	35
2.1 Evolutionary and developmental lineage of cell types.....	69
3.1 Mean birth weights among racialized groups in the US.....	95
3.2 Diagram of the mechanism schema for low birth weight births in African American women.....	104

Introduction

Philosophers take it that science is primarily concerned with providing explanations (e.g., how does the moon influence the tides? why is there diversity of life on earth?). Philosophy of science has offered competing accounts of what science is doing when it explains. Accounts of scientific explanation have variously argued that scientific explanation involves: *deducing* explananda-phenomena from laws of nature and initial conditions (Hempel and Oppenheim 1948); reducing the number of phenomena we need to appeal to by *unifying* seemingly disparate phenomena (Friedman 1974, Kitcher 1989); and, most prominently of late, identifying the *causes* of the phenomenon to be explained (Salmon 1984; Woodward 2003; Strevens 2008); among others.

Recently, some philosophers have proposed that, at least in the life sciences, explaining a phenomenon involves describing the *mechanism* that produces, underlies, or maintains that phenomenon. For instance, we explain the presence of an enzyme by pointing to the mechanism of protein synthesis, or we explain a rat's navigation of a maze by pointing to neural mechanisms of memory recall, and so on. Glennan (1996, 2017), Machamer, Darden, and Craver (2000), Bechtel and Abrahamsen (2005), Craver (2007), Craver and Darden (2013), among others, are influential statements of the mechanistic approach to scientific explanation. This approach, called the new mechanistic philosophy of science, has been fruitfully applied to disciplines ranging from molecular biology to economics and is particularly influential in philosophy of biology. On this approach, the *discovery of mechanisms* is both a description of how many sciences operate and is considered a standard for successful explanation.

New mechanists propose entities (or parts) and activities or interactions, as the explanatorily basic ontological components of mechanisms. That is, mechanisms are composed of entities and activities or interactions; mechanistic explanation consists of identifying these components and how they are organized to produce, underlie, or maintain a target phenomenon. Consider the mechanism of protein synthesis mentioned above; it includes material components, i.e., **entities**, such as **DNA**, **messenger RNA**, **ribosomes**, energy molecules such as **ATP** (adenosine triphosphate) and **GTP** (guanosine triphosphate), and **amino acids**, among others. It also includes causal components, i.e., *activities*, such as *transcription* (of DNA to pre-mRNA), *alternative slicing* (of pre-mRNA to mRNA), *translation* (mRNA to amino acid sequences), and *protein folding*. These entities and activities are organized so as to produce the explanandum-phenomenon: a protein. Mechanism schemas or models, through the description of these components and their organization, do the work of explanation.

Nonetheless, critics of the mechanistic approach to philosophy of science have charged that the ontological categories posited by new mechanists are philosophically unfounded and poorly demarcated. Specifically, critics charge that the new mechanists have not established a satisfactory standard on: 1) how to determine which *activities* perform the *causally productive* work in mechanisms, and 2) how to accurately identify and individuate the *parts* of mechanisms. My dissertation is motivated by these philosophical challenges. Chapters one and two of the dissertation answer each of these questions in turn by defending a novel account of activities (causes) and good parts. In the third chapter, I apply the mechanistic approach, and the accounts I develop, to questions of explanation in race and medicine. The dissertation is structured as follows.

In chapter one, I propose a novel, Hybrid Account of Activities (HAA) as productive difference-makers. The Hybrid Account incorporates difference-making into the predominantly production-based activities view of causation, thus bringing the two rival approaches to causation together. The union of production and difference-making in the Hybrid Account has a number of philosophical virtues. First, HAA enables the identification of causally productive activities and their distinction from causal influences that are not causally productive.

Second, unlike interventionist accounts of difference-making, HAA is able to distinguish *productive* difference-makers from mere background conditions. Finally, the Hybrid Account does all of this while preserving the attractive features of the new mechanists' approach, namely, their attention to the actual workings of science and scientific practice. The marriage of production and difference-making in the Hybrid Account forms the basis of the union of scientifically informed philosophy of science and philosophically informed scientific practice that is the ultimate aim of the new mechanists' approach.

On the Hybrid Account, activities are not one off, occurring once and never happening again. Rather, I defend an account of activities that provides a philosophical characterization of the *types* of activities biologists typically study¹. These activities are recurrent and produce the types of changes that are essential to the mechanisms biologists investigate. Activities are what produce the changes needed to account for the regular, productively continuous relations from start or set up to finish or termination conditions of mechanisms (including, sometimes, feedback loops and cycles). In sum, this chapter will develop an account of activities with two explanatory virtues that previous views of activities lack: (i) it integrates production and difference-making

¹ Given the role that regularities play in the account I develop, one-off mechanisms such as the extinction of the dinosaurs do not fit well into my account.

accounts in its characterization of activities, and consequently (ii) it distinguishes between causally productive activities and casual influences that are merely background conditions.

In chapter two, I defend a “carving standard” for parts. The new mechanistic philosophy of science makes what Kauffman (1971) called “articulation of parts” central to explanatory practice, at least in the life sciences. Bechtel and Richardson (1993) advocate decomposition and localization as a first step in discovering mechanisms. This calls out for a carving principle that justifies the standard parts that are included as components of mechanistic explanations. Franklin-Hall (2016) criticizes functional individuation of parts as insufficient to exclude gerrymandered parts such as “quarter-neurons” that are as causally efficacious as “standard” neurons. To address the insufficiency of causal efficacy as a carving principle, my account of parts states that the decomposition of a system that produces parts that do not fit into the *explanatory mosaic* is inferior to the one that does. And the world is such that many of the parts incorporated in mechanistic explanations fit into the explanatory mosaic. My appeal to the explanatory mosaic of science to ground good parthood adds to the condition that good parts (1) are material objects that (2) engage in activities and (3) have the (causal) properties that enable them to do so. As a consequence of the explanatory constraint I defend in this chapter, it is the richly textured explanatory mosaics produced by science (in a particular field at a given time) that grounds what we can say about how to carve parts at nature’s joints.

My account of good parts provides a carving principle that secures a valuable feature of good, standard parts. My account justifies the store of parts special sciences use in their explanatory work by grounding good parthood in explanatory mosaics of the sciences, an account of which I defend in this chapter. Roughly, the explanatory mosaics of science are how the foundational theories or mechanism schemas of a given science posit their functionally

integrated units “hang together” in a mutually explanatory web. I draw on Thagard’s (1989) influential work on explanatory coherence to develop my account. I argue that evolutionary and developmental lineages constitute an important explanatory mosaic in biology. These lineages ontologically give rise to good parts; and our scientific investigation of lineages enables us to judge which parts belong in the explanatory mosaic and are therefore good parts. I draw on lineages to show how certain forms of gerrymandering, such as “quarter-neurons”, can be ruled out. This both respects how scientists work while defending a philosophical principle that justifies their practice.

In chapter three, I apply the accounts of activities as productive difference-makers and of good parthood I develop in chapters one and two to the case of the use of race in epidemiology and biomedicine. The biomedical race debate concerns what role, if any, race should play in medicine. Much of the debate has revolved around the merits of biological racial realism and anti-realism, that is, on whether race is a meaningful biological category. I show how genetic-association studies lack the explanatory constraints to illuminate the causal structure of racialized differences in disease risk. I forward an alternative approach drawing on and expanding work in mechanistic explanation. Mechanism discovery offers a normatively and explanatorily attractive methodology to researching, diagnosing, and treating complex trait disorders. But it delivers the result that race will not be a medically relevant category in the manner the realist suggests.

To motivate my account, I examine the prevailing strategies used by biomedical researchers and philosophers of medicine to account for epidemiological disparities between racialized groups (Black, White, etc.) in the United States. The account I develop draws on a mechanistic approach to investigating these disparities that shows the explanatory balance between genetic, environmental, and other factors in accounting for them. I apply my approach

to the case of low birth weight (LBW) births among Black Americans. The approach I defend incorporates productive difference-making occurring at different sizes (from molecular to social) that are part of the productive continuity of the mechanism that produces the phenomenon. This approach centers the elucidation of all causally relevant factors, without wrongly focusing on a single factor (Reimers et al. 2019). Furthermore, it draws on the mutual compatibility of standards of causal relevance across the natural and social sciences (Baetu 2019). By showing how differential environmental exposure experienced by differently racialized groups (and sub-racial populations) interacts with genetic and developmental factors to generate epidemiological disparities, a mechanism discovery approach can resolve the tension between genetic and environmental approaches to epidemiological explanations.

Chapter 1:

The Hybrid Account of Activities

1.1 Introduction

This chapter defends a novel account of activities and how to identify and individuate them. Machamer, Darden, and Craver's [MDC] (2000) influential account of mechanisms as "entities and activities organized such that they are productive of regular changes from start or set up to finish or termination conditions" launched a burgeoning literature on mechanistic explanation (MDC 2000, 1). These New Mechanists characterized an account of explanation that is ubiquitous in at least sciences such as molecular biology and neuroscience. Their account has since been applied to a variety of sciences, and even extended to accounts of metaphysical explanation (Trogon 2018). Although this approach has become prominent in philosophy of science, a central component of the account has received relatively little philosophical attention, namely activities.

Activities are "producers of change" in the preferred characterization of MDC. This much is agreed upon by all proponents of the activities view of causation. Machamer (2004) and Bogen (2008) defended this view. However, both doubt there is any unifying characteristic that all activities share beyond being types of causes. Bogen (2008) endorses Anscombe's contention that it is a "brute fact" that activities have their effect and that "there is no informative general condition which discriminates causally productive activities from goings on which are not causally productive" (Bogen 2008, 113). Machamer (2004) is equally skeptical that there is a unifying feature of all activities. He writes of activities that "it is not clear that they all have any one thing in common or are similar in any significant way..." (Machamer 2004, 29).

This lack of unifying characteristics for activities has opened the activities view to the charge by critics that it is philosophically uninformative (Godfrey-Smith 2009, Franklin-Hall

2016). Godfrey-Smith (2009) has pointed out that without general conditions for identifying activities, we (especially scientists we might add) cannot meaningfully hypothesize about as yet undiscovered activities. To know what kinds of causal connections are produced by activities we must know what, in general, makes something a causally productive activity (Godfrey-Smith 2009, 334).

Glennan (2017), Kaiser (2018) and Krickel (2018) have addressed Godfrey-Smith's (2009) worry. Kaiser's (2018) and Krickel's (2018) "main features of activities" and Glennan's (2017) necessary and sufficient conditions provide what Machamer (2004) and Bogen (2008) had eschewed, the general conditions for being considered an activity. However, these accounts have serious limitations. Neither Kaiser (2018) nor Krickel (2018) provide a condition or feature of activities that mark them as distinct from causal influences that are not productive. Although Glennan has adapted activities as part of his characterization of mechanisms, he still proposes a mechanism-centered, and not activity-centered, conception of cause. This is because in order to account for how an activity is causal, Glennan requires that one go down to the mechanism that is responsible for the activity. In effect, the identification of activities as producers of change is merely an intermediate step to understanding the nature of causation. We must characterize the mechanism at a lower level to the phenomenon (the activity in this case) to account for its causal productivity.

Furthermore, activities-based views of causation have so far not accounted for what makes activities *productive*. Psillos (2004) characterizes the mechanistic and counterfactual approaches as two rival accounts of causation. Counterfactual accounts characterize causation as a form of dependence where the occurrence of the effect depends on the occurrence of the cause. According to the counterfactual approach causes *make a difference* to their effects. Psillos (2004)

discusses Woodward's interventionist account as his paradigmatic counterfactual approach. The mechanistic approach characterizes the relationship between cause and effect as one of *production*. Psillos (2004) argues that the production view taken by mechanisms is insufficient to account for the causality of mechanisms. Franklin-Hall (2016) further argues that the activities view fails to distinguish between causal production and causal influences that are irrelevant to production within the mechanism. Without a philosophically informative standard of what constitutes causal productivity, the activities view risks being a deflationary account of what scientists do instead of a philosophical account of causation.

In order to overcome these difficulties this chapter defends a novel account of activities as productive difference-makers of change, which I call the Hybrid Account of Activities (HAA). On the Hybrid Account, causally productive activities are what make a difference to changes between stages of a mechanism. Causally productive activities are *identified* by the particular kind of difference-making in which they engage (as specified by the Hybrid Account) and *individuated* by the types of changes they produce. The Hybrid Account marries the two dominant philosophical approaches to causation: production and difference-making. Previous activities views developed purely productive accounts of activities as causes. This opened the activities view to the "problem of irrelevant production." The problem, put simply, was the inability of production views to distinguish causal production from one stage of a mechanism to the next from the wider web of causal influence. The novel approach forwarded in this chapter is a union of production and difference-making that offers an attractive solution to this problem that preserves the virtues of the new mechanist approach.

On the Hybrid Account, activities are not one off, occurring once and never happening again². Rather, I defend an account of activities that provides a philosophical characterization of the *types* of activities biologists typically study. These recurrent activities produce types of changes from one stage of a mechanism to the next. Activities are what produce the changes needed to account for the regular, productively continuous relations from start or set up to finish or termination conditions of mechanisms (including, sometimes, feedback loops and cycles). The account I defend is not a *general* account of causation, or of an activity as such, but about activities as the causal components of *mechanisms*. Furthermore, the account I provide applies to *etiological* mechanisms with start or set up and finish or termination conditions. I do not discuss constitutive, part/whole mechanisms found in, for instance, Craver (2007). Finally, my account of activities has two explanatory virtues that previous views of activities lack: (i) it integrates production and difference-making accounts in its characterization of activities, and consequently (ii) it distinguishes between causally productive activities and casual influences that are merely background conditions.

The chapter goes as follows: Section 1.2 outlines the early activity view of causation defended by MDC (2000), Machamer (2004), and Bogen (2008), the challenge to this approach highlighted by Godfrey-Smith (2009) and recent accounts by Glennan (2017) and Kaiser (2018) that attempt to meet that challenge; section 1.3 discusses Franklin-Hall's and Psillos' criticism of the activities view of causation that it doesn't distinguish causally irrelevant side-effects from causally productive activities; section 1.4 briefly discusses Woodward (1997, 2003, 2010) and Strevens' (2004, 2008) difference-making approach and the latter's unification of difference-

² There may be other types of causes that serve explanatory purposes in the case of one-off mechanisms. And one-off mechanisms may have regular activities as a component. The Hybrid Account is about only one type of cause, namely activities, that appears in regular mechanisms.

making and production; in section 1.5, I propose and defend my Hybrid Account of Activities and develop the case of phosphorylation to highlight the virtues of the account; finally, section 1.6 concludes.

In what follows (section 1.2), I outline the activities view of causation. The section is better called the activities *views* of causation since there is a divergence among new mechanists who incorporate activities into their account of mechanisms on how they characterize activities. Nonetheless, the new mechanists discussed in the next section take activities to be the causal components of mechanisms. I discuss these views and their differences below and highlight their particular shortcomings.

1.2 The Activities View

The first activities view was proposed in Machamer, Darden, and Craver's (MDC) (2000) groundbreaking paper in what would become known as the new mechanistic philosophy of science. MDC characterize mechanisms as "entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions" (2000, 3). Mechanisms are not machines. That is, they are not "exclusively mechanical (push-pull) systems" (2000, 2). In order to be a mechanism, it has to be *active*. MDC drew on molecular biology and neuroscience in developing and motivating their account.

The MDC account of mechanisms is dualistic. Their ontology is composed of two irreducible and ontologically equal things: *entities* and *activities*. Entities are property-bearing objects that engage in activities. Activities are the producers of change. As MDC put it, activities are "types of causes. Terms like 'cause' and 'interact' are abstract terms that need to be specified with a type of activity and are often so specified in typical scientific discourse" (MDC

2000, 6). Some examples of activities include binding, folding, splicing, firing, and so on. These activities bring about changes within a mechanism such as activation of any enzyme (phosphorylation), silencing of a gene (methylation), modification of messenger RNA (splicing), and so on.

Entities and activities are organized within mechanisms to produce a phenomenon. “Entities often must be appropriately located, structured, and oriented, and the activities in which they engage must have a temporal order, rate, and duration” (2000, 3). Mechanisms have *regularity* in so far as they operate in more or less the same way under the same conditions. If the description of the mechanism is complete, then there must be *productive continuity* from the start or set-up to the end or termination condition. That is, there will be no gap in our knowledge of each step in the production of the phenomenon (2000, 3).

According to MDC, activities are not merely properties of entities, or entities’ capacities. If the universe were frozen in time, as if it were an insect encased in amber, all the entities along with their capacities and many of their properties would still be there, but there would be no activities. Without activities no mechanisms would be operating. Scientists are not merely after what things there are, but “how things work” (Craver 2007). Mechanistic explanation is well-suited to give us such an account. It is therefore unsurprising that scientists, especially life scientists, appeal to mechanisms in their explanatory work. Entities and activities are therefore irreducible components of mechanisms (Illari and Williamson 2013).

MDC’s (2000) characterization of mechanisms set off a burgeoning literature on mechanistic explanation. Machamer (2004) expanded on the activities view forwarded in MDC. Consistent with the characterization of mechanism and activities in MDC, Machamer (2004) defended the activities view of causation. Machamer (2004) claims that the criteria for the

identification of activities “are specific to kinds of activity that a group, at a time and in a discipline, takes to be fundamental in the sense that they do not feel any need to question their truth or usefulness” (Machamer 2004, 29). He adds: “It is not clear that they all have any one thing in common or are similar in any significant way, but neither commonality nor similarity are necessary conditions for an adequate category” (Machamer 2004, 29). While it may be true that we can make do with a category without necessary and sufficient conditions for membership in that category, Machamer’s (2004) contention that there is no unifying conception that applies to all activities leaves little justification for the heavy ontological weight MDC place on activities. As Glennan (2017) notes, if activities are what ground our causal claims, “there must be something about activities that makes them what they are” (Glennan 2017, 31).

Bogen (2008) defends an activities view of causation broadly similar to MDC (2000) and Machamer (2004). On Bogen’s account, that activities are causally productive is a further brute fact about activities. He writes that “an activity is causally productive by virtue of facts about the activity, the things that engage in it, and what results from them doing so” (Bogen 2008, 117). Bogen contends there is no “single, universally applicable criterion” that distinguishes causes from non-causes (Bogen 2008, 117). Both Machamer (2004) and Bogen (2008) claim activities are discovered and catalogued by scientists, but there is not much further that can be said to unify them.

Godfrey-Smith (2009) points to a problem with a “minimalist” approach to causation such as the one espoused by Machamer (2004) and Bogen (2008). Since the early activities view was skeptical of unifying conditions for activities, they are committed to a causal “minimalism” where “all we know is that some specific relations are grouped together by the word ‘cause’” (Godfrey-Smith 2009, 333). For proponents of the activity view, the grouping of these relations

is under the concept of ‘activity’. However, this “minimalist” approach poses a challenge to relying on scientific practice to identify which relations are activities. Since nothing further is said about the specific relations that count as activities, it is unclear what it is that enables scientists to know the relations they are investigating are in fact activities (or causes). As Godfrey-Smith (2009) puts it “if minimalism was true it would make no sense for a person in seventeenth-century England to wonder whether the Great Fire of London had somehow caused the end of a plague epidemic that had preceded it” since “we had not previously developed a more specific description for that type of connection and grouped it as causal” (Godfrey-Smith 2009, 334). Therefore, the “minimalist” approach taken by early proponents of the activities view undercuts the claim that it is ultimately the task of the scientist to discover activities. Without a conceptual account of what *constitute* activities, there is no framework with which to understand the discovery of *novel* activities. Recently, proponents of the activities view have begun to provide just those features in virtue of which something is an activity. Below, I discuss the accounts of Kaiser (2018) and Glennan (2017) (and their shortcomings) and provide my own response to these critiques.

Kaiser (2018) develops the activities view by identifying the main features shared by all activities (Table 1.1). Following MDC (2000), Kaiser (2018) takes “activity” to be how we further analyze “cause” within the context of mechanistic philosophy of science. That is, in order to account for causation and explain causal claims made by scientists, New Mechanists develop an account of activities. Activities “produce the changes” (MDC) in a mechanism in virtue of their “active” nature (Machamer 2004). Following Craver (2007), activities are the “causal components of mechanisms” (Kaiser 2018, 118). Activities are temporally extended, and therefore they belong to the metaphysical category of occurrents (118). This puts activities in a

different metaphysical category from entities (which are continuants), preserving the dualism of entities and activities initially proposed by MDC (2000). Kaiser claims activities, as the occurrent component of mechanisms, must be *actualized*. Activities are not dispositions or potentialities of entities. They are the actually occurring producers of change. Activities are always the activities of an entity or entities, that is without entities to engage in them there would be no activities. Additionally, without activities, entities would not be able to produce changes in themselves or in other entities. Finally, a description of a mechanism (mechanism schema) is complete if it shows the “productive continuity” of the mechanism from beginning to end.

Kaiser’s Main Features of Activities
K1 Activities are <i>temporally extended</i> (i.e., occurrents).
K2 Activities are <i>actualized</i> (rather than merely potential).
K3 Activities <i>produce change</i> (i.e., are types of causes).
K4 Activities require at least one <i>actively involved entity</i> .
K5 Activities have <i>unrestricted arity</i> (i.e., involve one to many entities).

Table 1.1: Kaiser’s (2018) list of the main features of activities (Kaiser 2018, 119 [italics in original])

Kaiser provides a helpful analysis of the new mechanist account of activities as causes. Indeed, the Hybrid Account I develop in section 1.5 assumes K1-K5 as constituting the core of the activities view of causation. However, if activities are to serve as the chief causal concept for new mechanists, the philosophical account needs to give the precise conditions in which K 1-5 obtain (Table 1.1). For instance, if activities are *producers* of change, how does this notion of “production” relate to causal relevance? How do new mechanists who propose an activities view

of causation distinguish causally productive activities from causal influences that are mere side effects or background conditions³? Kaiser’s main features do not answer this question.

Glennan (2017) has a view of activities that incorporates activities as part of his account of a “minimal mechanism” where “a mechanism for a phenomenon consists of entities (or parts) whose activities and interactions are organized so as to be responsible for the phenomenon” (Glennan 2017, 17). However, unlike earlier activity views, Glennan (2017) still relies on mechanism and not activity as his primary account of causation (more on this below). Glennan (2017) nonetheless provides what he labels the necessary and sufficient conditions for being an activity in response to Godfrey-Smith’s (2009) challenge. They are as follows:

Glennan’s Necessary and Sufficient Conditions for Being an Activity
G1 Activities require entities (parts, components) to act and be acted upon.
G2 Activities produce change in entities (parts, components) that act or are acted upon.
G3 Activities manifest the powers (capacities) of the entities involved in the activity.
G4 Activities are temporally extended processes.
G5 Most or all activities are mechanism-dependent.

Table 1.2: Glennan’s list of the necessary and sufficient conditions for being an activity (Glennan 2017, 31).

Glennan’s (2017) account goes further than Kaiser’s (2018) by grounding the claim that activities are producers of change (G2) in the claim that activities are mechanism-dependent (G5) (Table 1.2). Activities are productive because there is an underlying mechanism that is responsible for them. For Glennan (2017), “the mechanism-dependence of (all or most) production is a unifying feature of productive activities ...” (Glennan 2017, 33). What makes something causally productive, as opposed to a mere causal influence, is that there is an

³ This challenge is elaborated in section II. I consider the causal challenge of distinguishing production from causal influence to be the central problem facing the activities view.

underlying mechanism that connects a “causing event” with an “effect event” (Glennan 2017, 179).

Glennan (2017) goes further than previous activities views by providing a philosophical analysis of production. But his account suffers from a circularity. Glennan (2017) defines a minimal mechanism as “consist[ing] of entities (or parts) whose activities and interactions are organized so as to be responsible for the phenomenon” (Glennan 2017, 17). Activities are therefore already part of his account of mechanisms. Yet, Glennan (2017) goes on to account for the productivity of activities by claiming they are mechanism-dependent. Activities are characterized in terms of mechanisms, which is itself characterized partly in terms of activities⁴. Furthermore, Glennan’s (2017) proposed solution does not show us how to distinguish between the wider causal web of influence and the mechanisms he takes to be productive and therefore causal. The “problem of irrelevant production” is a major difficulty of production accounts of causation, especially since Glennan (2017) contends it is singular mechanisms in the world, and not their types, which ground the causal productivity of activities. Glennan’s (2017) contention that “difference-making is connected in the first instance not to completely concrete events, but rather to higher-level events, or, alternatively to abstract features of the concrete events” strains against his account’s singularist and actualist conception of mechanisms (Glennan 2017, 195).

⁴ Craver (2007, 89-90) had criticized Glennan (1996) for being circular since it accounted for causation through mechanisms, which are themselves organized causings. Glennan (2017, chapter 7) defends against this charge by claiming mechanisms at different levels are sufficiently different to escape this worry. But this isn’t a satisfactory response since Glennan (2017) explains the productivity of activities by appeal to underlying mechanisms. Does the activity of a mechanism at n-level get its productivity from the activities of the mechanism at n-1 level that underlie it or from the whole mechanism at n-1? If it is the former, we have a problem of regress (must I appeal to quantum mechanics to account for the productivity of the heart’s *pumping*?). If it is the latter, activities will lose all claim to being the causal concept on his account.

In summary, the early activities view of causation claimed that it is a further brute fact about activities that they are causal in just those ways scientists who study them claim they are. Godfrey-Smith's (2009) criticism of "minimalist" approaches to causation posed conceptual problems for discovery of novel activities. Glennan (2017) and Kaiser (2018) propose identifying conditions that apply to all activities. Although this advanced the activities view beyond the objection to the view's minimalism, their accounts nonetheless do not satisfactorily account for the causal productivity of activities. What is needed is a "feature" or "necessary and sufficient condition" of activities that accounts for their causal productivity. And this account should not reduce the causal productivity of activities to some other ontological category in order to remain a *distinctly* activities view of causation.

In the next section, I outline Psillos' (2004) Franklin-Hall's (2016) criticism of the activities view of causation. Namely, Psillos' charge that difference-making is needed for causal identification and Franklin-Hall's charge that activities views of causation fail to distinguish between causal production and causally irrelevant side-effects.

1.3 Activities and a Lack of Explanatory Constraints

Psillos (2004) has three criticisms of mechanistic approach, particularly of MDC (2000). First, MDC have not demonstrated that there is a non-epistemic reason to prefer an activities to a capacities view of causation. While we cannot know what capacities entities have without first observing what activities they engage in (as MDC claimed), it is nonetheless possible. Psillos argues, that it is the capacities of entities that do the causal and explanatory work. For instance, it is the capacities of salt and water that produce the dissolved salt effect, even though in order to

find out the capacities of salt and water we need to observe their myriad activities (Psillos 2004, 313).

Second, Psillos (2004) challenges MDC's claim that activities are necessary to account for "state transitions." Psillos (2004) counters that the capacities view is capable of accounting for state transitions as well as activities. Given the capacities of salt and water, there is nothing over and above the salt, water, their respective capacities, and the circumstances under which they are brought together, that is needed to account for the salt's dissolution in the water. The activity of "dissolving" only has, as stated above, an epistemic priority over the salt/water capacities. It is nonetheless unnecessary to account for the actual effect.

Finally, and most crucially, Psillos (2004) argues that MDC need counterfactuals to make the activities view work. Activities such as bonding, breaking, repelling, and so on can only be distinguished as causal if they build in a notion of difference-making. "X broke Y" is only a causal claim if a claim of the sort "had X not struck y, then y would not have broken" is true. Without counterfactuals, MDC is not able to extricate the causal (activities) wheat from the non-causal chaff. Glennan (1996) characterizes interactions as invariant relationships between parts of a mechanism that are stable over a range of counterfactual interventions. MDC leave activities as more ontologically basic and undertheorized. In both cases, Psillos (2004) argues, there is a causal-explanatory gap that is being filled, or needs to be filled, by a counterfactual view.

The final challenge lies at the core of Psillos' (2004) criticism of the new mechanist approach. Namely, there is an asymmetry between the counterfactual and the new mechanism views. The new mechanism approach depends on counterfactuals, while the counterfactual view stands alone. Even if there is nothing objectionable in the new mechanism approach, the counterfactual view is more basic and does most of the philosophical heavy lifting. However,

Psillos (2004) argues that mechanist considerations have a number of epistemic virtues. The process of discovering mechanisms, or “looking inside the box” to see how things work, is essential to having a full understanding of the causal relations at work (Psillos 2004, 317).

In a similar vein, Franklin-Hall (2016) charges that activities do not form an explanatorily apt account of causation. More specifically, she argues that the activities view commits a *causation error*. In order to motivate her challenge, Franklin-Hall proposes a *mechanistic model*, or a mechanism schema in the Craver and Darden (2013) terminology, of a neuron’s release of “neurotransmitters at its axon terminal when its dendrites are exposed to neurotransmitters, and not otherwise” (Franklin-Hall 2016, 44). The Standard Model explains this capacity by relating organized macromolecular parts such as axons, dendrites, ion channels and gradients, with “dynamic principles” which causally relate input to outputs. But she proposes a non-standard alternative model which is identical to the standard model,

except that it appeals to two alternative dynamic principles, one relating neurotransmitter exposure and membrane vibration, and a second relating vibration and any later event genuinely relevant to neurotransmitter release, for example, the entry of calcium into the axon terminal. With these principles and others, such an alternative model might bridge inputs and outputs, stating first that neurotransmitter exposure is followed by membrane vibration, itself followed by cellular calcium entry, eventuating finally in neurotransmitter release (Franklin-Hall 2016, 46).

However, the “alternative model” commits a causation error. The vibration of a membrane is not part of the productive continuity of neuron depolarization. Yet, Franklin-Hall (2016) contends, it still appeals to “organized parts changing according to dynamic principles” (Franklin-Hall 2016, 46). The incorporation of membrane vibration as a part of the mechanism of neuron depolarization is unscientific. It is experimentally known that neuronal membrane

vibrations are the result of the wave of ions rushing through the neuron and do not actually produce the depolarization itself (Carlen et al. 1982).

The new mechanists who defended the early activities view do forward a principle that eliminates the “alternative model” in favor of the standard explanation. MDC write of mechanism components, including activities, that the “the components [are those] that are accepted as relatively fundamental or taken to be unproblematic for the purposes of a given scientist, research group, or field” (MDC 2000, 13). Therefore, since no competent scientist would cite membrane vibration as part of the productive continuity of neurotransmitter release, we have reason to pick the standard over the alternative model.

However, Franklin-Hall (2016) argues looking to the sciences to identify and individuate activities is an unsatisfactory way to block the causation error since this ends up being too deflationary. Scientists often identify activities, but we need a philosophically informative account of how they come to correctly identify the parts of mechanisms, including activities. Otherwise, mechanistic explanation is purely a science reporting exercise (Franklin-Hall 2016, 47). Furthermore, taking the Machamer (2004) and Bogen (2008) position of leaving the causal productivity of activities as an unexplained brute fact “is completely opaque”, leaving no room for the view to say why the membrane vibration model is incorrect beyond that it is not what competent scientists would claim in an explanation (Franklin-Hall 2016, 53).

Franklin-Hall’s (2016) criticism hits upon an important limitation of earlier activities views of causation. A desideratum of any account of causation is that it provides the conditions that enable the selection of just some elements of the vast causal web of influence as cause(s) of an event. This is no less the case for the activities view. A mechanism operates within a much larger causal nexus, perhaps including cosmic influences (e.g., cosmic radiation). Mechanism

schemas are often composed of multiple parts organized such that changes from one stage to the next are accounted for by activities. This requires, both at the scientific and philosophical level, a condition that identifies the activities, out of a panoply of occurrences happening simultaneously, as the ones productive of the change. The puzzle of how to account for this selective feature of causation is taken up by difference-making (or relevance) accounts of causation. Difference-making is often contrasted with production. I argue that activities are both producers *and* difference-makers (specifically, their production is a special kind of difference-making). The vibration is not a productive activity because it never makes a difference to the occurrence of the next stage of the mechanism. Unifying difference-making and production in the activities view offers an attractive account that addresses the criticisms discussed above. In what follows, to motivate my account of activities, I first briefly outline difference-making approaches to causation and Strevens' (2004, 2008) unification of these concepts.

1.4 Difference-Making and Causal Production

Difference-making approaches to causation and explanation variously propose a way to separate the causal wheat from the non-causal chaff. These approaches to explanation have become dominant in philosophy of science with Woodward (1997, 2003, 2010), Strevens (2008), Waters (2007), and Weslake (2009) forwarding influential accounts. Difference-making approaches to causation are attractive because they accord with both everyday causal attributions and scientific practice. From an incalculably wide range of causal influence we pick out just one or a few elements that we identify as genuinely causal. Difference-making approaches characterize this as largely involving ignoring or abstracting away from the causal web of influence all those details that do not make a difference, i.e., are irrelevant, to the effect or explanandum obtaining.

An influential difference-making account of causation is Woodward's (1997, 2003, 2010) interventionist or manipulability view of causation. For Woodward, to say X explains Y is to say that X causes Y, where X and Y are variables that can have at least two values. Explanations answer what-if-things-had-been-different questions (*w*-questions) that a difference-making approach is well-suited to answer. Woodward takes the relationship between cause and effect to be a form of counterfactual dependence. However, the problem of relevance plagued earlier counterfactual accounts of causation. The joint effects of a cause may counterfactually depend on one another. Nevertheless, they do not provide information about causation. For instance, a barometric reading and a storm counterfactually depend on each other in so far as whenever the barometer displays a reading indicating low air pressure there is a storm and vice versa. However, the barometric display does not cause the storm, neither does the storm cause the barometric display. Rather, both the barometric reading and the storm are the effects of a common cause, namely, low air pressure.

To avoid the problem of irrelevant counterfactuals, Woodward provides an account of which counterfactuals count as causal. According to Woodward, "the counterfactuals that matter for explanation are counterfactuals the antecedents of which are made true by a special sort of exogenous causal process that I call an intervention" (Woodward 1997, s29).

Woodward defines interventions as follows:

An intervention on X with respect to Y is an idealized experimental manipulation of X which causes a change in Y that is of such a character that any change in Y occurs only through this change in X and not in any other way (Woodward 2010, 290).

Interventions in the account above should be interpreted *heuristically*. That human intervention is not possible does not detract from the view. Suitably characterized natural

phenomena may also count as interventions (e.g., a lightning strike). Furthermore, the appeal to intervention, which is itself a causal term, is not viciously circular as Woodward is not providing a reductive account of causation but elucidating how certain causal relationships can be explained in terms of others. This accords with the fact that experiments are widely taken to provide causal information. In other words, “in order to test some causal claims we must assume the truth of others” (Woodward 1997, s31).

Nevertheless, Woodward’s interventionist account of difference-making is too broad to capture causally productive activities while excluding causal influences scientists normally take to be background conditions. Craver and Kaplan (2020) claim they have not yet found a satisfactory way to identify the “differences that make a difference” beyond what can be established by scientific inquiry. They write that “sometimes, we draw a line between foreground and background conditions (*though background conditions often make a big difference*) (Craver and Kaplan 2020, 26 [italics added]). However, they have “little of general interest to say about [the] considerations” which lead us to assign to some difference-makers the status of background conditions and to others the status of cause. It is this weakness of Woodward’s account I plan to address with my Hybrid Account of Activities.

Before I provide my account, I briefly consider another prominent approach to difference-making forwarded by Strevens (2004, 2008). His approach unifies difference-making and production in a way I make use of in my account.

Strevens (2004, 2008) defends a causal account of explanation that incorporates an essential role for difference-making. In Strevens’ approach, which he terms the “kairitic account”, difference-making plays the role of picking out “the explanatorily relevant parts of any causal network” (Strevens 2004, 158). For Strevens (2008), the relationship between cause and

effect is one of entailment. An effect follows from its cause. A causal model's set of the premises represent the causal influences that jointly entail an effect. The difference-makers are represented by those premises whose elimination from that set would lead to a failure to entail their effect (Strevens 2008, 70-83).

To briefly see how this works, take an effect *E*. First, from within the causal web in which *E* is embedded, take "a set of actual initial conditions and laws" that was sufficient for the causal production of *E*. The representation of this set is a "veridical deterministic causal model" (Strevens 2004, 162). Second, remove from the model anything, a condition or law, without a role in the causal production of *E*. The initial conditions and laws that remain after the eliminative procedure are difference-makers. The model that remains is a "explanatory kernel" for *E* (Strevens 2008, 118). Strevens in effect ends the dichotomy between difference-making (or relevance) and production. For Strevens difference-making is a feature of the premises in a causal model in virtue of which they are able to causally produce their effect.

The account of activities I advance in the next section shares with Strevens (2008) this unifying approach to these two concepts of causation. I show how activities are causally productive by showing that they are special kinds of difference-makers. However, unlike Strevens (2008) I do not take the relationship between cause and effect (in my case activities and the change they produce) to be one of derivational entailment. I draw on Woodward's (2010) intervention account of difference-making. But unlike Woodward, I add a further constraint on difference-making to identify just those difference-makers that are *productive* of the next stage of a mechanism and not merely background conditions. I now turn to my Hybrid Account of Activities.

1.5.1 A Hybrid Account of Activities as Productive Difference-Makers

Activities produce change, but this fact does not tell us how to identify and individuate activities. Moreover, on the MDC account, activities are “productive of regular changes” (MDC, 2000, 2). To this production account I add difference-making: activities are the actual difference-makers to the occurrence of the changes from one stage of a mechanism to the next. This could be the transfer of energy, the transmission of information (for instance the “precise determination of sequence” in the DNA-RNA-protein schema), among others. Activities produce change by making a difference to the realization of the next stage in a mechanism. For instance, in the well-characterized mechanism of protein synthesis, the entity RNA polymerase *transcribes* DNA to messenger RNA (mRNA). The activity of *transcription* produces the next stage of the mechanism (in this case the entity mRNA). I argue activities are individuated by eliminating (i) occurrences that do not make a difference to a change and (ii) those difference-makers that do not make a difference in the highest proportion of the scenarios that that change occurs. What remain are the causally productive activities. What follows is the Hybrid Account of Activities:

(HAA) Φ -ing is a causally productive activity of a change ψ in the next stage of a regular mechanism if:

- i) Φ -ing makes a difference to the occurrence of a change ψ in at least one of the scenarios in which ψ occurs and,
- ii) In the highest proportion of the scenarios that Φ -ing occurs, the change ψ occurs
- iii) Φ -ing makes a difference to a narrow range of types of changes ψ_{i-n}

The first condition (i) incorporates Woodward’s interventionist difference-making into an account of activities. This move blocks concerns that the activities view does not distinguish

between causal production and irrelevant causal influence. On my account, we can remove putative causal influences from a mechanism schema if they never make a difference to the production of a type of change. If intervening on a variable of a causal influence does not change the variable of a given change ψ , then it is not a causally productive activity. We can now discount the membrane vibration in the alternative model proposed by Franklin-Hall (2016) as a causally productive activity. The membrane vibration is not an activity, and hence does not belong in the mechanism schema or model, because it does not make a difference to the occurrence of the next stage of the mechanism. An idealized intervention on the membrane vibration would not prevent the occurrence of the next stage of the neurotransmitter release mechanism.

While the first condition answers the Franklin-Hall (2016) challenge, there are still many difference-makers within a mechanism that are typically not considered among the causally productive activities. Many causal influences biologists often take to be background conditions, such as temperature and pH, fulfill condition (i) of my account. That is, background conditions are often difference-makers and intervening on them will make a difference to the occurrence of the effect. The second condition answers the problem with distinguishing background conditions from causally productive activities that Craver and Kaplan (2020) highlighted. Condition (ii) of my account differentiates these difference-making background conditions from causally productive activities.

If we take the set of difference-makers to the occurrence of a change (the next stage of a mechanism), the causally productive activity is the member of that set that, in the highest proportion of the scenarios in which it occurs, the change (next stage of a mechanism) occurs. Background conditions fail HAA (ii) because they occur in many more of the scenarios where

the change does not occur than the causally productive activity. For instance, if we take changing from an inactive to an active form of an enzyme (i.e., a regulatory protein) as a type of change, the activity *heating at 37 Celsius* makes a difference to that type of change. However, this heating activity occurs in many scenarios where that type of change does not occur. It therefore is a background condition and not a causally productive activity. Of course, the change can fail to occur because the causally productive activity has been blocked. For example, catalytic activities that activate an enzyme can be blocked by an inhibitor. In this case, the causally productive activity will fail to occur. But it still holds that in the highest proportion of the cases the causally productive activity occurs, the change occurs. This is how the Hybrid Account identifies which difference-maker is *productive* of a type of change. HAA (ii) captures the appeal of production views of causation, which make the effect expectable⁵ from the cause since the causes produce the effect.

HAA (iii) is the specificity condition; it further distinguishes between activities and background conditions. Causally productive activities satisfy HAA (iii) while background conditions do not. The domain over which this condition quantifies is given by the types of changes to which putative causal influences (background conditions and activities) make a difference. To take one example, the internal body temperature of mammals makes a difference to the circulation of blood, hormonal control mechanisms, enzyme activity, the nervous system, and so on. The range of the types of changes to which temperature is a difference-maker is immense. However, activities such as pumping, phosphorylating, and transcribing make a

⁵ Bhogal (2020) defends an account of explanatory goodness whose first dimension, precision, holds that “explanations are better if in more of (that is, a higher proportion of) the physically possible worlds where the explanans is true, the explanans explains the explanandum” (Bhogal 2020, 18). He notes that the intuitive force behind “precision” is the classic philosophical idea that explanations should make the explanandum expectable. While HAA is not an account of explanatory goodness, HAA (ii) is a condition that is meant to capture the same ideal for the activities view of causation.

difference to a *narrow* range of changes. Pumping moves material (blood, ions, molecules) into or out of an enclosed space, phosphorylating produces ATP and activation of an enzyme, and transcribing produces an mRNA molecule. The effect-changes to which activities make a difference are few, while causal influences that count as background conditions, such as the gravitational influence of celestial bodies, temperature, pressure, pH, have a wide range of effect-changes to which they make a difference.

It is also important to keep in mind that HAA is consistent with more than one activity producing a type of change. HAA (ii) merely claims that if Φ -ing is to count as a causally productive activity, then in the highest proportion of the scenarios Φ -ing occurs within a mechanism, the change ψ occurs. It does not follow from this that in all the scenarios that the change ψ occurs, Φ -ing occurs. For instance, methylation (adding a methyl group) may inactivate a given gene X in the cases it occurs. Yet, gene X can be rendered inactive by other activities. As long as those activities also fulfill HAA's conditions, they can be considered causally productive activities. This is an attractive feature of the Hybrid Account since it accords with scientific work on activities. There may be several types of activities that produce the same type of change in different mechanisms.

A mechanism is *regular* if there are multiple tokens of the same mechanism type that operate in the way described by the mechanism schema (i.e., it is not a one-off mechanism like the big bang). The scenarios that form the domain over which HAA quantifies are the tokens of (types of) mechanisms. Mechanism tokens are of the same type if 1) they are composed of similar entities and activities organized in a similar way and 2) they produce the same phenomenon. Conditions 1 and 2 are jointly necessary and sufficient for mechanism tokens to be of the same type. For instance, the F-Type ATP synthase mechanism has similar entities and activities

organized to produce the same phenomenon (ATP), tokens of this mechanism are found in bacterial plasma membranes, mitochondrial inner membranes, and chloroplast thylakoids. ATP synthase's role in producing the basic energy units of cells makes it one of the most important constituent mechanisms in living organisms, and "the similar overall structure and the high amino acid sequence homology indicate that the mechanism [...] is the same in all organisms" (Deckers-Hebestreit and Altendorf 1996, 791).

It is possible for different mechanism types to produce the same phenomenon (e.g., digital vs. analog clock producing time keeping). And there exists variation among tokens of the same mechanism type that nonetheless produce tokens of the same phenomenon type. The notion of similarity used in the characterization is not one of identity. One way to cash out similarity is *interchangeability*. If two mechanism tokens, *a* and *b*, are of the same type, every component of *a* can be replaced by the components of *b*, and vice versa, and still produce the same type of phenomenon. Two tokens are sufficiently divergent to be different types if any of their components are not interchangeable.

In biological systems, the interchangeability condition draws on evolutionary, developmental, and property-based considerations as a basis for claims of interchangeability. Components of mechanisms have a particular evolutionary and developmental history, and entities have activity-enabling properties. Just as a mechanic can safely assume all the parts manufactured according to the same specifications can be used as replacement for one another, evolutionary and developmental history, along with knowledge of specific activity-enabling properties, can be the basis for an interchangeability claim in biological systems. Interchangeability accounts for both variation among mechanism tokens and forms the basis for the claim that variant tokens are similar enough to be considered the same type.

The Hybrid Account relies on what Craver and Darden (2013) call the “store of types of entities and activities” (Craver and Darden 2013, 75). These are the types that scientists have characterized and investigated, and which they draw on in the work of proposing and discovering mechanisms. I add to that conception the idea of a type of change. Since many biological mechanism types share the same stages, a type of change from one stage of a mechanism to another is part of that store. The types of changes found in the store enable the *individuation* of causally productive activities. A type of change is potentially found in multiple mechanisms (e.g., activation of an enzyme) and produced by multiple types of activities. It is not only phosphorylation, but also hydroxylation and glycosylation, among other activities, that produce this common type of change (enzyme activation). This individuation procedure, where types of activities are indexed to a type of change, can be applied to the whole “store of types.” Answering Godfrey-Smith’s (2009) worry (discussed above), HAA provides the conditions for identifying activities. HAA can therefore potentially facilitate the discovery of new activities if, for instance, biologists find a type of change that is not produced by the activities already in the store.

These three features of activities highlighted by HAA, that they are difference-makers, that their occurrence, more than any other causal relationship, results in the effect-change, and that they are specific, marks activities as a special type of cause that are well-suited for the life sciences. Not all causes must fulfill these conditions in order to be considered causes. But within the mechanistic explanatory project, it is *activities* as specific types of causes that play the causal-explanatory role as producers of change within a mechanism.

In what follows, I’ll explicate my account by showing how it enables identification and individuation of causal productive activities in a molecular and a physiological case.

To see how the Hybrid Account of Activities works, take for instance the activity “phosphorylating.” Phosphorylation is a one of the most important activities in living organisms. It is part of mechanisms ranging from gene regulation to cancer pathogenesis. If we take this activity to be the phenomenon in need of explaining, we find that there is a complex mechanism that underlies it, with multiple steps involving bond-formation, spatial and temporal organization, and ATP (an energy source in cells) metabolism. Phosphorylation produces the changes it does by adding a phosphoryl group to a specific site of a protein. The addition of this phosphoryl group changes the shape of the protein to which it binds. The protein will then have the structure necessary to carry out its activity (Pearlman et. al 2011). Below, I illustrate phosphorylation activity with the case of glycogen phosphorylase.

Glycogen metabolism is a homeostatic mechanism that maintains stable blood glucose levels by synthesizing and degrading glycogen, particularly in the liver. Glycogen is the form carbohydrates take during storage. Liver glycogen is synthesized in response to an increase in the insulin to glucagon ratio as a result of increased glucose concentration in the blood after a meal. Glycogen is degraded between meals and releases as glucose in the blood to maintain blood glucose levels. The synthesis mechanism is mediated by (liver) glycogen synthases.

Human liver glycogen metabolism has four stages (activities in italics):

- i) Phosphorylase kinase produces active glycogen phosphorylase by *phosphorylation*.
- ii) Glycogen phosphorylase produces glucose 1-phosphate by *phosphorolysis* of glycogen.
- iii) Phosphoglucomutase produces glucose 6-phosphate by *phosphorylating* glucose 1-phosphate.
- iv) Glucose 6-phosphatase produces glucose by *hydrolyzing* glucose 6-phosphate.

Stage (i) involves change from inactive (GP_b) to active (GP_a) form of the glycogen phosphorylase enzyme by the *phosphorylation* (adding phosphate group) of the serine 14 region of the GP_b form (discussed in detail below). In stage (ii) the active glycogen phosphorylase

engages in *phosphorolysis*, which involves using an energetic phosphate group to cleave bonds, to release a glucose 1-phosphate from a glycogen chain (glycogen is typically a chain of 1000 glucose molecules). In stage (iii) phosphoglucomutase *phosphorylates* glucose 1-phosphate to glucose 6-phosphate by transferring a phosphate group from the 1 to 6 position. Finally, in stage (iv), glucose 6-phosphatase *hydrolyzes* glucose 6-phosphate releasing a free glucose molecule that can enter the blood and transport to tissue in need of free glucose (Berg et al. 2002; Adeva-Andany et al. 2016).

As discussed above, (liver) glycogen phosphorylases initiate glycogen degradation (stage i). Importantly for our purposes, “only the phosphorylated form of liver phosphorylase (GPa) is catalytically active. Interconversion between GPa and GPb (unphosphorylated) is dependent on the activities of phosphorylase kinase and of phosphorylase phosphatase” (Agius 2015, 33 [underline added]). In other words, phosphorylation is the activity that produces the change from the inactive liver phosphorylase (GPb) to the active form (GPa) of the enzyme, and in all the scenarios that liver phosphorylases are phosphorylated, they are activated.

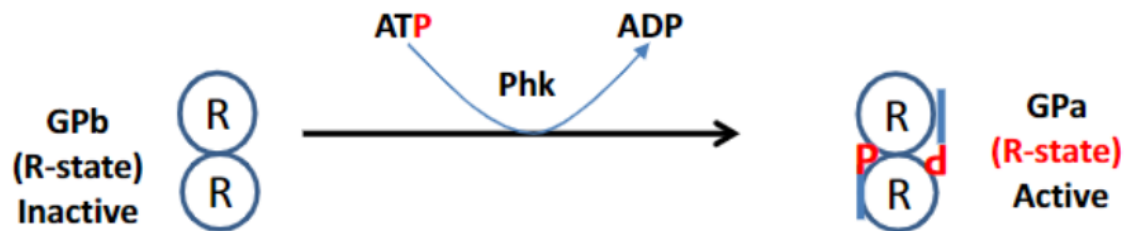


Figure 1.1: Phosphorylation of inactive glycogen phosphorylase (GPb) by phosphorylase kinase (Phk) turns it into the active form of the enzyme (GPa). Phosphorylation activity involves taking a phosphate group from another substrate (in this case ATP) and adding it to a binding site of a protein (in this case serine-14 of glycogen phosphorylase) [image adapted from Agius 2015, 36].

Specifically, the serine-14 portion of glycogen phosphorylase enzyme has a binding site for phosphates. Phosphorylase kinases are entities that carry out the phosphorylation activity

(Figure 1.1). Therefore, we have a type of change, inactive to active form of glycogen phosphorylase, produced by an activity (phosphorylation) that makes a difference to the occurrence of this change in all the scenarios this occurs.

As Figure 1.1 illustrates, the presence of ATP (an energy molecule that supplies the phosphate) is crucial for the phosphorylase kinase's ability to engage in the phosphorylation activity. Intervening to eliminate ATP would prevent the phosphorylation and subsequently the activation of glycogen phosphorylase. Nonetheless, scientists identify the phosphorylation carried out by phosphorylase kinase as the causally productive activity and the ATP as a background condition. HAA provides philosophical justification for this assignment. ATP, while a difference-maker, is present in many of the scenarios where glycogen phosphorylase is in its inactive (GP_b) form without the change to the active form (GP_a) occurring (in fact, ATPs are ubiquitous in living cells). That is, ATP is present in many more of the scenarios where the change (from inactive to active) does not occur than phosphorylation. ATP therefore fails to make the change expectable in the way the causally productive activity of phosphorylation does. Phosphorylation, however, is the difference-maker that makes a difference to, and therefore produces, the activation of glycogen phosphorylase in the highest proportion of the scenarios it occurs. ATP activities are not what drive the mechanism of glycogen metabolism forward. Their presence merely facilitates what does, the activity of phosphorylation.

The Hybrid Account shows that the mechanism-dependence condition (G5 in Table 1.2) defended by Glennan (2017) is unnecessary. While *explaining* how an activity works requires going down to the mechanism that underlies it, it is not necessary to appeal to this mechanism in order to identify the causally productive activities. What matters for activities qua causes is that they are productive difference-makers to a type of change. Any variation between the

mechanisms that underlie an activity type, such as phosphorylation, are not relevant to its causal productivity. For instance, in eukaryotes and prokaryotes phosphorylation proceeds differently. Different kinds of enzyme complexes, encoded by different genes, engage in phosphorylation. Then why is it that these are all considered the same *type* of activity, even though the underlying mechanisms are different? The reason is that they are all *productive difference-makers of a certain type*. It is phosphorylation, adding a phosphoryl group, that productively makes a difference to the occurrence of the next stage in multiple mechanisms ranging from photosynthesis to cell signaling. The causal productivity of phosphorylation as an activity lies in the fact that it makes a difference just in the way outlined by HAA, not in the underlying mechanism that produced it. Distinguishing between *explaining* activities, which requires describing the underlying mechanism, and *identifying* them as causally productive, which is what the Hybrid Account does, preserves the activities view of causation and avoids the problem of circularity or regress engendered by Glennan's (2017) G5 (see fn. 2).

To take another example, one of the earliest enzymes where this type of change (from inactive to active) was characterized was tyrosine hydroxylase. This enzyme is important in many neural mechanisms. Active tyrosine hydroxylase is essential in the synthesis of dopamine, an important neurotransmitter in neuro regulation. The activation of tyrosine hydroxylase is a type of change that serves as the start condition of the dopamine synthesis mechanism (Daubner et al. 2011). In their groundbreaking work Joh et al. (1978) characterize how this type of change comes about:

...the pool of native tyrosine hydroxylase is composed of a mixture of enzyme molecules in both active and probably inactive forms, that the active form is phosphorylated, and that *phosphorylation produces an active form of the enzyme at the expense of an inactive one* (Joh et al. 1978, 4744 [italics added])

This characterization of their discovery illustrates my HAA account. They experimentally identify phosphorylation as the activity that makes a difference in this type of change (from inactive to active tyrosine hydroxylase). They conclude this because they have discovered what Darden and Craver (2002) call “activity signatures.” Phosphoryl groups, a signature of the phosphorylation activity, was found on all active tyrosine hydroxylases.

Let us take a further, physiological, example to illustrate the Hybrid Account’s ability to block Franklin-Hall’s (2016) criticisms. Consider the circulatory system of mammals, an important entity is the heart, and its crucial activity pumping. The fact that blood circulates was an important discovery in the history of anatomy. Within the circulation mechanism, the movement of deoxygenated blood to the lungs and oxygenated blood from the lungs through the rest of the body is enabled by the heart’s pumping activity. First, it counts as an activity on my account because (i) it makes a difference to the change Ψ (movement of blood) and (ii) in the highest proportion of cases pumping occurs, the type of change Ψ occurs. In fact, pumping is the only way organisms with a circulatory system are able to move blood. Even arthropods, which have open circulatory systems where the blood flows freely, have muscles in their abdomen that have pumping-enabling properties. Invertebrate animals such as sponges have no blood and move nutrients through a different activity, diffusion.

The pumping activity of the heart illustrates how the Hybrid Account avoids Franklin-Hall’s (2016) charge of *causation error* against mechanistic explanation. For instance, one can propose a non-standard mechanism schema for circulation that has the stage contraction of the myocardium, followed by a lub-dub sound, followed by blood leaving the aorta (a major blood vessel). This, of course, contains a non-standard mechanism module (namely the lub-dub sound)

that although related by “dynamic principles” to the next stage, nonetheless is not considered part of the production of circulation. Franklin-Hall (2016) charges that new mechanists have not worked out a way to separate non-standard from the standard explanations like the one I gave above. Simply asserting that this non-standard mechanism is not productive of the circulation phenomenon, even though it is correctly describing (at least in part) the behavior of the parts of the standard mechanism, is not philosophically informative.

The Hybrid Account can, however, meet Franklin-Hall’s (2016) challenge. The reason the heart’s lub-dub sound is not an activity is that it fails condition (i) of my account. It doesn’t make a difference to the blood’s exiting (or entering) the heart in *any* of the scenarios that type of change occurs in the mechanism of circulation. The lub-dub sound of the heart is a detail that can be abstracted away in our characterization of the mechanism of circulation, while the heart’s pumping cannot. In fact, in cases of Still’s murmurs, benign heart murmurs that develop in some children, the heart does not make a lub-dub sound at all, even as it continues to produce the next stage in the circulation mechanism. This is more so the case with artificial ECMO (extracorporeal membrane oxygenation) machines that move blood in patients undergoing heart procedures. The machine may make beeping and buzzing noises but not lub-dubs, and what allows it to perform the heart’s function in that medical context is its pumping activity. Therefore, one can conclude that lub-dubs do not make a difference to the movement of blood. It is pumping, and not lub-dub sounds, that make a difference and in all the scenarios that pumping occurs, movement of blood occurs.

The Hybrid Account of Activities also addresses Psillos’ (2004) challenges. First, by incorporating difference-making along with production in the account of activities, it grants to Psillos (2004) that activities have to be characterized as difference-makers in order to distinguish

them from the non-causal. However, I challenge Psillos' (2004) claim that the virtues of the mechanistic approach are purely epistemic. On the contrary, the Hybrid Account is able to draw a scientifically significant distinction between background conditions and causally productive activities that counterfactual dependence accounts do not. In order to see how, I contrast HAA with a previous attempt to provide principles to narrow causal influences to specific causes forwarded by Woodward (2010).

1.5.2 HAA as a Superior Approach to Causal Specificity

Woodward (2010) expands on his influential interventionist account of causation (and therefore scientific explanation) to elucidate three features of causal relationships that are of particular importance in biology, namely stability, proportionality, and specificity. Stability is a feature of causal relationships that hold under a wide range of background conditions. Proportionality is a feature of causal relationships where changes in the cause are mirrored proportionally by changes in the effect. Specificity, which draws on Lewis' concept of influence, is a feature of causal relationships that approach the one-cause-one effect ideal.

Woodward (2010) provides a minimal account of causation:

(M) X causes Y if and only if there are background circumstances B such that if some (single) intervention that changes the value of X (and no other variable) were to occur in B , then Y or the probability distribution of Y would change. (Woodward 2010, 290).

Woodward incorporates the notion of background conditions into his definition of M. He writes, "*background circumstances* are circumstances that are not explicitly represented in the X–Y relationship" (290).

Causal relationships are stable, according to Woodward (2010), if they continue to obtain under a wide range of background conditions. Stability comes in degrees. Causal relationships are more or less stable depending on how wide or narrow the range of background conditions under which they obtain are. Woodward (2010) is not trying to distinguish between background conditions and “stable” causes. Rather, *given* a large set of possible background conditions, causal relationships are more/or less stable depending on whether they are invariant (the relationship continues to hold) under a wider or narrower subset of those background circumstances (292).

Whether a cause is proximate or distal often has bearing on whether a causal relationship is more or less stable. Distal causes tend to be (although they aren't always) less stable than proximate causes. To take Lewis' (1986) example, writing recommendation letter L caused X's grandchildren to exist is relatively unstable. Only under a narrow range of background circumstances does L lead to X's grandchildren. If X's would be spouse attended a different program, or they never met romantically, and so on, the writing of L would not cause X's grandchildren to exist. These small changes are able to derail the link between L and X's grandchildren because L is a *distal* cause. On the other hand, the writing of L caused X to be accepted into the program is relatively more stable since L is a *proximate* cause of that effect.

Crucially for the biological sciences, stability is related to whether it is easier or harder to alter causal relationships, a question that is important in biomedicine and evolutionary theory. Natural selection has more opportunities to alter less stable relationships (295-6). For instance, changes in the less stable genetic regulation of splicing can produce huge phenotypic variation even while the DNA to pre-mRNA relationship (which is highly stable) remains unchanged. This generates the variation on which selection acts.

Activities, according to HAA, are stable under Woodward's (2010) characterization. Activities are proximate causes since they produce changes from one stage of the mechanism to the next. Therefore, as a general rule, activities will have stability in the sense outlined by Woodward (2010). HAA has built into the account a desired characteristic of biologically significant causal relationships. That is, identifying biologically significant causal relationships will often involve identifying activities. HAA achieves this while also providing an account to distinguish background conditions from causally productive activities, an account that is missing from Woodward's notion of stability.

Woodward (2010) also discusses the causal notion of specificity, which is ubiquitous in the biological sciences (Woodward 2010, 292). Woodward (2010) develops specificity by (i) drawing on Lewis' notion of influence and (ii) the one-to-one conception of causation. Specificity is often invoked to account for the privileged role DNA plays in explaining the development of various phenotypes. Woodward (2010) argues that the best way to conceive of specificity, and the special causal role of DNA, is as a form of fine-grained influence. According to influence, "C will influence E to the extent that by varying the state of C and its time and place of occurrence, we can modulate the state of E in a fine-grained way" (Woodward 2010, 305). Woodward (2010) illustrates influence with an analogy to a dial on a radio. Moving the dial on a radio will change the stations in a fine-grained way and a given position of the dial is associated with a particular radio station (307). Causal relationships that lack specificity on the other hand are more switch-like. Whether the radio is plugged in or not for instance is not causally specific since it does not have fine grained influence on the operations of the radio. It is in this sense that DNA is distinct from the other "cellular machinery" in causing various phenotypes. Changes in RNA polymerases, ribosomes, energy molecules (ATP, GTP) does not

have fine grained influence over the product. However, in the case of DNA “there are many possible states of the DNA sequence and many (although not all) variations in this sequence are systematically associated with different possible corresponding states of the linear sequences of the mRNA molecules and of the proteins synthesized” (306).

Specificity may be a candidate for a criterion to distinguish between background conditions and specific causes. Background conditions, which tend to be switch-like, are those causal relationships that are less specific. The more specific a cause is, the less we identify it as a background condition and the more we identify it as a biologically significant cause or activity. However, there are problems with this approach. Many biologically significant causal relationships, which we would not consider mere background circumstances, have switch-like roles. This means on Woodward’s account they will not have causal specificity. Activities such as phosphorylating, methylating, and acetylating, do not admit of dial-like modulations; they either occur or do not and their effect is determinative (to activate an enzyme, to silence a gene, etc.). Of course, many activities do have dial-like, as opposed to switch-like, influence in the changes they produce (e.g., diffusion). However, this distinction does not constitute the fundamental cleavage between background conditions and causes (activities); rather, it cuts across them.

Woodward (2010) considers another standard for specificity, which sees specificity as approximating a one cause-one effect relationship. This notion of specificity draws on epidemiology where identifying a one-to-one relationship between a cause and a disease-effect is the gold standard of epidemiological research. For instance, cases where being exposed to a particular toxin causes one disease, or a specific gene variant causes one disorder, and so on. Of course, most diseases have a many-cause many-effect causal relationship and do not fit the one-

to-one standard. Woodward (2010) amends this notion by marking it not as a criterion of causation, but as a biologically significant type of causal relationship. In this sense of specificity, “—*C* will be a more (rather than less) specific cause (in the one to one sense) to the extent that it causes only a few different kinds of effects within a pre-specified range” (311). For example, enzyme activity is more specific if it interacts with only a narrow range of substrates and produces only limited number of effects. Activities on the Hybrid Account are relatively specific in this second sense. Although not all activities produce only one effect, they do produce “only a few different kinds of effects within a pre-specified range.”

Woodward’s (2010) three causal notions, particularly stability and specificity, identify important conceptions of biologically significant causation. However, they do not address a core conceptual question, namely, why are some causally relevant relationships background conditions while others are biologically significant causes? Is this merely a consequence of the explanatory priorities of scientists? The Hybrid Account provides a non-interest dependent answer to this question. Additionally, activities are stable and specific. HAA therefore builds the notions of biological significance Woodward (2010) outlines.

1.5.3 HAA and the Right Balance Between Generality and Specificity

The examples given above to illustrate HAA show how it provides an attractive principle for identifying activities and distinguishing them from irrelevant causal side-effects and background conditions. However, it might still be the case that HAA is too *broad* a principle and identifies as a causally productive activity something scientists would not endorse. HAA might also be too *narrow* and exclude as an activity, or identify as a background condition, a specific cause for a given phenomenon. I discuss a potential case of each below. First, I turn to the role of

concentration in molecular mechanisms and whether HAA (ii) might end up being too broad a principle.

Enterobacteria phage λ (henceforth Lambda) is a bacterial virus that infects *E. coli*. Lambda has two lifecycles: lysogenic and lytic. During the lysogenic cycle Lambda inserts itself into the *E. coli* genome and remains dormant, thereby preventing replication of the phage within the bacterium. During the lytic cycle Lambda produces multiple copies of itself until the bacterium bursts (lysis). Lambda encodes two proteins that regulate which cycle the phage is in: Cro and cI repressors. Whether Lambda is in its lysogenic or lytic stage is controlled by six regulatory sites on its DNA. If these sites are bonded to a cI repressor, the Lambda will remain in its lysogenic (latent) stage. The bonding of cI to the six controlling sites is frequently disrupted. Nevertheless, a high concentration of cI is maintained making it likely that these sites are bonded to a new cI repressor even though heavy churn continually knocks off cI repressors from their binding sites. However, there is still a low concentration of the Cro molecule in the lysogenic cycle. One of the six sites maintains the high concentration of cI repressors by upregulating the production of cI molecules. If this site is bonded with a Cro molecule, the production of cI will decline, reducing the concentration of cI molecules. In the low concentration of cI environment the latency produced by cI repressors binding to the six regulatory sites will become unlikely. Without cI repressing phage production, Lambda phages enter the lytic cycle, reproduce continually, and burst out of the *E. coli*.

The activities involved in the two cycles of Lambda raise a number of philosophical puzzles in causation and explanation. When it comes to maintaining the lysogenic cycle the activity of bonding, the fact that these six sites are bonded to cI molecules, is what maintains the lysogenic cycle. However, since bonding is not continuous, many cI molecules are knocked off of the

binding site, but under conditions of high concentration of cI molecules they are likely to be replaced by other cI molecules. There is therefore a dynamic and stochastic component to this activity. The cI bonding itself is what maintains lysogeny, and the Cro bonding to the site that upregulates cI production itself is what produces the change to the lytic cycle. However, the *high concentration* of cI molecules is necessary to maintain the lysogenic cycle.

There are three possibilities for the role played by high concentration of cI in maintaining the lysogenic cycle:

1) The high concentration of cI molecules is a background condition for the lysogenic cycle, not the causally productive activity. We can take *bonding* of cI to the six controlling sites as the activity that produces the lysogenic stage. The high concentration of cI is merely an enabling or background condition for the bonding activity. The problem for this view is that it is inconsistent with HAA. The *high* concentration of cI molecules makes a difference to the phage being in the lysogenic cycle (HAA (i)), in the highest proportion of the scenarios there is high concentration of cI, the phage is in lysogenic cycle (HAA (ii)), and high concentration of cI molecules is specific to a narrow range of types of changes (HAA (iii)). Since it fulfills all three conditions high concentration of cI would be a candidate for a causally productive activity and not a background condition, which takes us to the second possibility discussed below.

2) The second possible role of high concentration of cI is as a causally productive activity. As shown in 1 above, high concentration fulfills all three conditions of HAA. However, this outcome seems unsatisfactory. As Strevens (forthcoming) notes, it seems wrong to attribute a causal role to aggregates when the vast majority of the components of that aggregate do nothing at all. Most of the cI molecules in the “concentration” don’t do anything, in my terms do not engage in any activity. It therefore seems wrong to attribute causal productivity to the

concentration itself. Indeed, if anything the high concentration of cI molecules is a property of the population of the entities in question and not an activity. Perhaps then the role of high concentration of cI molecules in maintaining the lysogenic cycle is as an activity enabling property.

3) A third possibility for the role of high concentration of cI molecules is as an activity enabling property. But the high concentration of cI does not really enable bonding itself. The relevant activity is bonding, and the activity enabling property that is crucial is the one that belongs to individual cI molecules that enables them to “fit” with the binding sites on the phage DNA. whether or not the concentration of cI was high, any cI molecule can bond with the regulatory sites in so far as the size, shape, and orientation of cI molecules is compatible with the DNA sites. All the high concentration of cI does is make it *probable* that the cI will bond to the regulatory sites. But this probability cannot be said to be in the world as activities or properties are.

As shown above, it seems none of the three possibilities adequately captures the role of high concentration of cI molecules in maintaining the lysogenic cycle of Lambda. The high concentration is not merely a background condition as it fulfills HAA (ii) and (iii). However, it seems unsatisfactory to consider concentration a causally productive activity⁶ since it is bonding and not the concentration itself that produces the lysogenic cycle. It is also not an activity enabling property. The relevant property that enables *bonding* belongs to the individual cI molecules and not the concentration.

⁶ Nathan (forthcoming) defends the view that concentration can be considered robustly causal. I do not expand on the view here but merely note that this view is a departure from the mechanistic literature.

Another worry about HAA is that it might prove to be too narrow and exclude specific causes that are scientifically significant. In particular, HAA (iii) excludes from identification as activities all the causal influences that make a difference to a wide range of types of changes. This captures causal specificity in many significant scientific explanations. For instance, HAA (iii) identifies temperature, pH, gravitation, and so on as background conditions. Indeed, these causal influences are background conditions in a great number of mechanisms. However, they can also play the role of specific causes in some mechanisms. Gravitational pulling, for instance, is causally productive in pulley mechanisms to name just one example.

The two worries mentioned above show the difficulties of seeking necessary and sufficient conditions for activity-hood. While having necessary and sufficient conditions allows for easy identification and individuation of activities, the choice of conditions might result in some things being left out or included that do not accord with ordinary scientific or philosophical views. However, HAA may nonetheless strike the *best* balance of principles that ends up satisfying (most) philosophical desiderata while according with scientific practice.

1.6 Conclusion

This chapter proposes a novel, Hybrid Account of Activities (HAA) as productive difference-makers. The Hybrid Account incorporates difference-making into the activities view of causation. The union of production and difference-making in the Hybrid Account has a number of philosophical virtues. First, HAA enables the identification of causally productive activities and their distinction from causal influences that are not causally productive. Second, unlike interventionist accounts of difference-making, HAA is able to frequently distinguish *productive* difference-makers from mere background conditions. Finally, the Hybrid Account

does all of this while preserving the attractive features of the new mechanists' approach, namely, their attention to the actual workings of science and scientific practice. The marriage of production and difference-making in the Hybrid Account forms the basis of the union of scientifically informed philosophy of science and philosophically informed scientific practice that is the ultimate aim of the new mechanists' approach.

Chapter 2:

Good Parts and the Explanatory Mosaic

2.1 Introduction

In this chapter, I provide a principle that identifies and individuates parts⁷ of mechanisms. As we have seen in earlier chapters, MDC's (2000) influential account of mechanisms posited both activities, discussed in the previous chapter, and entities (or parts) discussed in this chapter (MDC 2000, 1). Mechanism schemas or models would not be explanatory unless they described how these parts and activities are organized so as to produce, maintain, or underlie explananda-phenomena. However, this raises the question of why any particular decomposition of parts should be privileged as *the* parts (along with activities) that produce a given phenomenon. The issue of carving has come up in recent criticism of the new mechanistic philosophy of science, with Franklin-Hall (2016) highlighting the absence of an adequate carving standard for parts as a philosophical limitation of the new mechanist approach. This chapter addresses the carving problem by defending a novel account of good parthood that philosophically justifies the standard parts used by scientists in their explanatory work.

The chapter goes as follows: in section 2.2, I outline what I call the carving problem as proposed by Franklin-Hall (2016) and discuss possible solutions in the existing literature. In section 2.3, I propose and defend my account of good parthood and show its relation to Thagard's (1989) account of explanatory coherence. I defend an account of the explanatory mosaics of science that grounds good parthood. In section 2.4, I show how good parts in some fields of biology have robust explanatory relations due to the evolutionary and developmental lineages to which they belong while gerrymandered parts do not. Finally, section 2.5 concludes.

⁷ Following Craver (2007) and Franklin-Hall (2016), I use "parts" to refer to the physical components of mechanisms while activities are the causal components of mechanisms. Parts in my usage corresponds to the "entities" of mechanistic literature (see MDC 2000; Craver and Darden 2013; Glennan 2017). The previous chapter addressed how activities are identified and individuated.

2.2 The Carving Problem

Franklin-Hall (2016) charges that new mechanists have not provided an adequate carving standard that picks out good parts and blocks “gerrymandered” parts. The paradigm mechanism schemas from molecular biology and neuroscience proposed by new mechanists such as MDC (2000) are composed of complex parts such as neurons, cells, DNA molecules, and membranes. Franklin-Hall (2016) argues the “carving standard” used by MDC (2000) and other new mechanists does not adequately block “gerrymandered” parts from their mechanism schemas or models. For instance, Franklin-Hall (2016) proposes an alternative decomposition of the mechanism of neurotransmitter release that cites “quarter-neurons,” “which are large chunks of biomass” that are still connected in the manner described by the standard schema of neurotransmitter release. The “quarter-neuron” has the properties of neurons that enable its causally efficacy (Franklin-Hall 2016, 46). “Quarter-neurons” can therefore potentially engage in causally productive activities. These “gerrymandered” parts can be proposed as entities or parts of mechanisms, even though no scientist considers them to be “good” parts. The upshot of the carving challenge is that it is possible, for any given phenomenon, to propose alternative mechanisms composed of parts carved differently that preserve the properties and causal relationships necessary to produce or constitute the phenomenon, while never being part of the scientific store of types of entities or activities.

The recent new mechanist literature has extended the discussion of parts. In the previous chapter, I discussed Glennan’s (2017) and Kaiser’s (2018) accounts of activities and their main features. Glennan (2017) and Kaiser (2018) also provide accounts of parts and their main

features (see Tables 2.1 and 2.2). However, the main features proposed by their respective accounts are inadequate to address the carving problem and block the gerrymandering objection.

Glennan (2017) contends that the concept of entities (henceforth parts) is “thin” and nothing substantive can be said of entities *as such*. For Glennan (2017), it is particular kinds of entities that can be “detected, individuated, and classified” (Glennan 2017, 34). Nevertheless, Glennan (2017) provides a set of necessary conditions for parthood that he argues captures one of the two core components of mechanisms (Table 2.1).

Glennan’s Necessary Conditions of Entity-hood
(G1) Entities are what engage in activities and interactions.
(G2) Entities have locations in space and are stable bearers of causal powers (or capacities) over time.
(G3) The causal powers of entities are what allow them to engage in activities and thereby produce change.
(G4) Most or all entities are systems composed of parts and most or all of the powers of entities will be mechanism-dependent.

Table 2.1: Glennan’s list of the necessary conditions of being an entity (Glennan 2017, 34).

(G1) and (G2) are elements of the standard view about parts among new mechanists. (G3) incorporates the concept of “causal powers” into the new mechanist account. Although Glennan (2017) doesn’t provide a detailed account of causal powers, as he puts it in his illustrative example—faculty having the “causal power” to change the curriculum and thereby producing change in students—suggests causal powers on his account are akin to activity-enabling properties (discussed below). When it comes to (G4), Glennan (2017) argues that an entity is mechanism-dependent insofar as “entity S is a system with parts X, and the entity’s capacities

and activities are constitutively dependent upon the activities and interactions of those parts” (Glennan 2017, 34f). That is, a part is characterized by its sub-parts and their causal interactions.

Regardless of whether G1-G4 are in fact necessary conditions for parthood, they fail to answer the carving problem and block the gerrymandering objection. Gerrymandered parts engage in activities, have locations in space, bear causal powers as stably as the parts from which they are gerrymandered, and are composed of sub-parts which themselves causally interact with one another. Consequently, Glennan’s (2017) conditions cannot prevent gerrymandered parts such as quarter-neurons from counting as good parts of mechanisms.

Kaiser’s Main Features of Entities
(K1) Entities are <i>material objects</i> (i.e., continuants).
(K2) Entities are <i>bearers of properties</i> , which allow them to engage in specific activities.
(K3) As components of mechanisms, entities necessarily <i>engage in activities</i> (at least once during the mechanism).
(K4) Entities can be <i>actively or passively involved</i> in activities.
(K5) Entities are relatively <i>stable and robust</i> .

Table 2.2: Kaiser’s (2018) list of the main features of entities (Kaiser 2018, 119 [italics in original])

Kaiser’s (2018) list of features of parts fares no better than Glennan’s (2017) when it comes to the carving problem. Gerrymandered parts such as “quarter-neurons” straightforwardly fulfill K1-K4⁸. Gerrymandered parts are material objects, and they bear properties (indeed activity-enabling properties), they engage in activities, either passively or actively. K5 is a

⁸ Following Franklin-Hall (2016), I am assuming “quarter-neurons” have the properties necessary to effectuate neurotransmitter release. In any case, Kaiser’s list does not escape the problem of gerrymandered parts.

potentially discriminating principle qua gerrymandered parts. Kaiser (2018) characterizes robustness and stability partly as a function of existing “in a variety of different mechanisms in different organisms and they retain their properties also if studied in isolation” (Kaiser 2018, 118). This is a step in the right direction. However, there is a risk of circularity. If we ground the goodness of parts in the fact that they figure in “different mechanisms in different organisms” and characterize mechanisms as composed of good parts (and activities), we have provided a circular characterization of mechanisms and good parts. The goodness of parts must be (at least) partly grounded in non-mechanistic factors in order to avoid this circularity.

In addition to the conditions and features discussed by Glennan (2017) and Kaiser (2018), below I consider candidate principles for identifying and individuating good parts that block gerrymandering, namely: parsimony, activity-enabling, boundary, and technological identification. I discuss each in turn.

2.2.1 Parsimony

One approach to blocking the gerrymandering problem is to appeal to the principle of parsimony. Simply put, the principle of parsimony states that, *ceteris paribus*, the simplest of competing explanations is the best. When adopted to the case of parts, *ceteris paribus*⁹, parsimony would favor simpler ways of carving up nature to ones that are more complicated. Parsimony would therefore prefer neurons to quarter-neurons since the former do not multiply the number of parts beyond necessity.

However, the principle of parsimony cannot successfully block the gerrymandering objection. Gerrymandering of parts can creatively be deployed not only to generate

⁹ For a view of *ceteris paribus* laws see Pietroski and Rey (1995).

“unnecessary” splitting (quarter-neuron, half-neuron) but also lumping (neuron+). We can conceivably “carve” a neuron to include the synaptic cleft and generate neuron+. This part is gerrymandered in so far as the synapse is not considered a part of a neuron. Yet neuron+ is more parsimonious than positing neurons *and* synapses. In order to overcome the gerrymandering objection, we need a carving principle that shows why our parts are neither multiplying nor *combining* beyond necessity.

2.2.2 Parts as Activity-Enabling

Parts engage in activities. Darden (2002) defends the view that parts engage in activities because they possess *activity-enabling properties* (Darden 2002, S363). Darden (2002) outlines the role of activity-enabling properties in a discussion of how these properties (in this case the polarly charged bases of DNA) enable prediction of the next stage of the mechanism and its associated activities. Nearly all accounts in the new mechanist literature, including my own, consider activity-enabling a necessary condition for being a good part. For instance, the myelin sheath is a lipid-based insulating layer that surrounds axons (and other cells of the nervous system) and enables the rapid transmission of the action potential along the neuron. Myelination is necessary for normal motor, sensory, and cognitive function. In the case of the neurodegenerative disease multiple sclerosis (MS), the immune system attacks and damages the myelin, eliminating its insulating activity (Steinman 2001). This results in the degradation of motor function in patients with MS. Without the insulating activity of myelin, action potentials propagate across neurons at considerably reduced speed thereby limiting sensory and motor coordination. Possessing a myelin sheath is the activity-enabling property of rapid propagation of action potential.

While activity-enabling is a necessary condition for good parthood, it is not a sufficient condition. As we have seen with the case of “quarter-neurons”, it is possible to gerrymander parts such that their activity-enabling properties are preserved even though they are carved non-standardly. “Quarter-neurons” do have myelin sheaths. Franklin-Hall (2016) noted causal efficacy can carve parts in multiple ways, each preserving the properties that enable causal production. Causal efficacy fails to provide a standard that excludes gerrymandered parts (“quarter-neurons”). Kauffman (1971) placed the weight of identifying the admissible decomposition of systems into various parts solely on whether or not the parts are causally relevant to the production of the explanandum system. He writes:

A successful articulation of parts explanation distinguishes between irrelevant causal consequences and important causal consequences of a part, thus the explanation not only accounts for the behavior of the whole, it supplies a view of what it is that the parts themselves shall be seen as doing from among the indefinitely many possible things each part might be taken to be doing (264-5).

While causal efficacy within a mechanism is an important condition of good parthood, it is not the sole criterion by which parts should be individuated for the reasons I have highlighted above.

2.2.3 Parts Delineated by Boundaries

Historically, parts, especially anatomical parts, were often identified and individuated before their real activities were discovered. In *On the Usefulness of Parts of the Body*, Galen writes that “the liver is the source of the veins and the principal instrument of sanguification” (Galen 1968). While Galen was wrong about the activity of the liver, the anatomical structure he identified as the liver, and indeed the anatomists that came centuries before him so identified, is the selfsame structure we still call the liver. The identification of anatomical structures is

simplified by the existence of “natural boundaries.” The liver is a single, bounded entity that is distinct from its surrounding. The heart, lungs, cells, and other structures similarly have natural boundaries.

Nevertheless, natural boundaries are not always what delineate good parts. Consider the case of nodes of Ranvier on neurons. These are un-myelinated nodes on axons where the cell membrane of axons is exposed to the extracellular environment. In such a case is the myelin sheath or the axon cell membrane the “natural boundary” between the axon and the environment? Depending on the answer different ways of carving neurons according to which natural boundary is selected can be supported. Cases such as muscles, which are composed of bundles of fibers, are “untidily” delineated, to take another example.

2.2.4 Technological Identification of Parts

Craver (2007) argues that good parts are “robust” in a view drawing on Wimsatt (1981). That is, good parts should “be detectable with a variety of causally and theoretically independent devices” (Craver 2007, 132). Advances in technology such as x-ray crystallography, fractal analysis, electron microscopy and so on have greatly enhanced our ability to individuate parts beyond simple human visual identification through the use of a multiplicity of devices implementing different methodologies (Wimsatt 1981; Sedivy et al. 1999; Craver 2007).

Nevertheless, a technological solution to identification of good parts runs into problems. Devices designed to detect biological and chemical structures do not always agree on the details of the target structure (Mitchell 2020). The fact that no single methodology or technological process can provide a definitive delineation of parts reveals a general problem for this approach. Namely, there are theoretical commitments we bring to the building and interpreting of the

output of these devices in order to interpret the results. There isn't always a "given" part that is the output of all these devices without the interpretive theories we bring to them. As such, we cannot rely on a pre-theoretical or purely technical solution to the carving problem.

In what follows, I outline and defend my preferred account of good parthood.

2.3 A Novel Account of Good Parthood

Below I propose an account of good parthood. Although I do not argue this is a set of necessary and sufficient conditions for good parts, it captures two core desiderata: it distinguishes good from gerrymandered parts and it philosophically vindicates the scientific store of entities. My account goes as follows:

P is a good part of a mechanism *M* if,

- (i) *P* has activity-enabling properties,
- (ii) *P* engages in at least one activity in *M*, and
- (iii) *P* is an element of the explanatory mosaic

(i) and (ii) capture commonly defended conditions on good parthood in the mechanism literature, including in Glennan (2017) and Kaiser (2018). Good parts are components of mechanisms and their role is to engage in activities. In order to engage in activities, they need activity-enabling properties.

However, unlike Glennan (2017) and Kaiser (2018), I do not solely stress the *material* aspect of parts, i.e., that they are material objects located in space. The carving problem suggests a need to move from a *material* standard to an *explanatory* standard for good parthood. Stressing the material basis of parts is not sufficient to exclude gerrymandered parts such as quartet-

neurons, which are also material, real objects. We need a standard that excludes such parts from consideration in scientifically privileged mechanisms. To do so, I add condition (iii) to my account of good parthood.

I call (iii) the *connectedness* condition on good parthood. My account brings together activity-enabling and a new notion of robust explanatory relations that grounds the goodness of parts. Being an element (or component) of an explanatory mosaic distinguishes good from gerrymandered parts. Although I appeal in the *connectedness* condition to an *explanatory* mosaic, I do not defend an account of explanation in this chapter. For our purposes, a minimal account of explanation will do and to that end I make use of Woodward's interventionist account of explanation discussed in chapter two. While I defended the Hybrid Account of Activities as the causal component of mechanisms in chapter two, I draw on interventionist causation to characterize good parthood. This is because "explanatory relation" is much weaker than "productive activity." *Activities* produce the change from one stage of the mechanism to the next. They are the component responsible for productive continuity in a mechanism. An *explanatory relation* (between A and B) merely indicates that intervention in a variable of A directly changes a variable of B. I draw on the weaker explanatory relation, rather than productive activity, for two reasons.

First, relying on explanatory relations allows my account to avoid circularity. Explanatory relations obtain not only within mechanisms but also within lineages and lineage-based explanations that are not mechanistic (see 4.1 below). As such, my account can draw on non-mechanistic explanations to ground good parthood. Second, experimental intervention plays a role in identifying good parts. For instance, scientists use knock-out experiments, where either naturally occurring or experimentally manipulated mutants are used to investigate the role of

gene(s) in the development of a given part. Bruce and Patel (2020) use this approach to identify the origin of insect wings (Bruce and Patel 2020). Therefore, experimental relations play a role in determining good parthood independent of fulfilling the higher bar of being productive activities.

The minimal, interventionist notion of explanation suffices to ground what my account means by explanatory relations. It is minimal because a great many causal influences count as explanatory on this account. However, the explanatory relations that matter are those that are invariant across a range of interventions. Hence my account limits relevance for determination of good parthood to *robust* explanatory relations.

The notion of *robustness* used in (i) has to do with the stability of the explanatory relations between the properties (not Wimsatt (1981) and Craver's (2007) sense of robustness as detectible by multiple techniques). That is, they must hold across a wide range of interventions. For instance, there is a robust explanatory relation between the *structure* of a protein (e.g., hemoglobin) and its *function* (e.g., binding and transporting oxygen in the bloodstream)¹⁰. The relation between these two properties is so deep that "structure determines function" is a frequently invoked mantra in molecular biology. Robust explanatory relations form what I call an explanatory mosaic. The explanatory mosaic is nothing over and above the (explanatory) relations between parts and properties. We can think of good parts as a dense area of this mosaic where a pattern of explanatory relations obtains. Below, I show how my account provides a carving principle that vindicates the standard parts identified by science. The example cases are drawn from biology, but the connectedness condition that this discussion is meant to highlight

¹⁰ It is the *relation* between structure and function that is robust, not the structure of the protein itself. Proteins often change their structure in the course of their varied activities.

can be generalized to other special sciences that engage in mechanism discovery and mechanistic explanation.

My account draws on Thagard's (1989) influential discussion of explanatory coherence. Thagard (1989) defends a theory of explanatory coherence as an approach to explanatory inference which he implemented in a connectionist algorithm called ECHO. His theory is meant to secure the truth and acceptability of scientific hypotheses. For Thagard, explanatory coherence is primarily a relation that holds between propositions. It follows from this, Thagard claims, that coherence is also a property of a set of related propositions and single propositions as well. Two propositions cohere if there is some explanatory relation between them. As Thagard puts it "in ordinary language, to cohere is to hold together, and explanatory coherence is a holding together because of explanatory relations" (Thagard 1989, 436). To say that two propositions, P and Q, cohere is to say that P is part of the explanation of Q; Q is part of the explanation of P; P and Q are together part of the explanation of some R; and P and Q are analogous in the explanations they give of some R and S. The more explanatory relations propositions have, the higher their acceptability. As Thagard puts it, "those hypotheses that participate in many explanations will be much more coherent with the evidence, and with each other, and will therefore be harder to reject" (Thagard 1989, 465). In addition, explanatory breadth is an important desideratum in scientific reasoning. That is, all else equal, we should prefer hypotheses that explain more than their alternatives.

One of Thagard's illustrative cases is Darwin's theory of natural selection. Thagard claims two of Darwin's main hypotheses, "DH2: organic beings undergo natural selection" and "DH3: species of organic beings have evolved", cohere. These hypotheses-propositions both explain each other and are together part of the explanation of many other important hypotheses

and evidence that make up Darwin's theory. The alternative hypothesis that Darwin contends with, that species were individually created by God, is rejected because it fails to cohere with the evidence Darwin adduces. Thagard shows how the evidence and hypothesis statements drawn from Darwin can be shown to cohere in ECHO while the creationist hypothesis does not.

Thagard's account has been criticized on several fronts including that it does not provide an account of explanation, and is therefore empty and unrelated to scientific acceptability, and that it fails to capture the value of prediction in scientific reasoning, among others. I will not here defend Thagard's account. I merely note that my account of good parts and the explanatory mosaic differs from Thagard's account of explanatory coherence on three important grounds. First, I do not provide an account of the truth or acceptability of a scientific hypothesis (or any other proposition). Second, my account does not place equal weight on all explanatory relations. Third, I draw on a minimal, causal approach to explanation to account for what is *explanatory* in explanatory relations (discussed above).

However, like Thagard I take it that scientific fields place a high value in their explanatory work to how proposed entities (and hypothesis) fit into the overall explanatory picture. I make use of this fact to ground my account of good parthood. Gerrymandered parts on my account share with Thagard's rejectable propositions that they are isolated. They fail to fit into the mutually explanatory relations as good parts do, lacking the latter's explanatory breadth. Below, I discuss how my account of explanatory mosaics, and the biological lineages at its core in biology, can ground good parthood.

2.3.1 The Explanatory Mosaic

As Craver and Darden (2013) argue, scientific explanation involves “implicitly call[ing] up the host of explanatory concepts, the store of accepted entities, activities, and organizational structures that people in the field are licensed to use in constructing a mechanism schema” with each scientific field having its own “store” (Craver and Darden 2013, 67). Special sciences have a specialized store of types of entities and activities that they draw on for explanatory work (Craver and Darden 2013). As Sider puts it “a special science, quite generally, attempts to explain a certain target domain of facts by means of certain theoretical concepts” (Sider 2013, 28). The scientific store of parts and activities, generated by respective special sciences, plays a central role on my account in the determination of good parts. I argue further that the scientific store constitutes what I call an *explanatory mosaic*.

The explanatory mosaic is distinct from Craver’s (2007) “mosaic unity of science”, which he develops as an account of how science is integrated across levels. It is also distinct from the “Humean mosaic” outlined by Lewis. On Lewis’ account, “all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another” (Lewis 1986, ix). That is, the world is made up of physical spacetime points (or point-like objects) and the spatiotemporal relations between them. On my account, however, the mosaic is composed of robust explanatory relations. It makes no metaphysical claim about what the ultimate or fundamental reality is.

The idea behind the explanatory mosaic adds *organization* to Craver and Darden’s (2013) “store.” The store of entities and activities is not a random assortment. Scientists do not randomly place items into the store. Rather entities in the store fit together in a particular mutually supporting configuration that I am calling an explanatory mosaic. To take an example

in biology, *evolutionary and developmental* lineages organize the store of biological entities in what is called the tree or web of life. Elements of the tree include whole species, a “Tree of Cells” in the case of monocellular organisms, and lineages of genes or individual parts such as eyes or feathers, among others.

Of course, as science progresses the entities it takes to be part of the store change as well. To briefly take an example, phlogiston was, for instance, superseded by oxygen as the dominant explanatory entity in what we now call oxidation-reduction reactions. Phlogiston theory posited explanatory^x relations between combustion, respiration, “calcination”, and the entity phlogiston. The superscript x indicates that there was no minimally causal explanatory relation between phlogiston and these host of phenomena. The explanatory^x relations posited by phlogiston theorists did not in fact obtain. The explanatory inadequacies of this theory in comparison to oxygen theory led to its eventual rejection in favor of oxygen theory (Thagard 1989).

In the next section, I explore a core feature of biology that forms an important element of the explanatory mosaic of the life sciences, namely, lineages.

2.4 Lineages as Sources of Explanatory Relations in Biology

Evolutionary and developmental lineages form one explanatory mosaic in biology. They play a role in explaining dynamic evolutionary processes (Calcott 2009). The scientific understanding of lineages is of course subject to revision. Whatever shape the explanatory mosaic takes, whether a tree or a web or some combination of the two, the basic “units” are bound together through robust explanatory relations and are therefore liable to be the good parts. The persistence of lineage explanations in biology is a further testament to their explanatory value.

In order to motivate my account of good parts, and therefore principle for individuating good parts, I first briefly discuss a common approach to the identification and individuation of parts in biology, namely functional individuation, and show its shortcomings. For those with a selected effects view of function, functional individuation is guided by what the putative part is taken to be selected for and the specific trajectory in evolutionary history it took. According to this view, “selected function is the means by which biologists *define* the parts and processes of the organisms they study” (Griffiths 2006, 3). Craver (2001) and Griffiths (2006), on the other hand, appeal to *causal* function, which is the causal role of a part in biological systems, to individuate biological parts. As discussed in section 2, causal-functional individuation (i.e., maintaining the causal efficacy of the part) can yield non-standard parts and is therefore not a sufficient carving standard. And selected function of parts might exclude biological domains where selection history is not relevant to identification of a part (Love 2007).

However, I argue functional individuation is not a sufficient parthood standard and fails to exclude gerrymandering of parts. On my account, both causal and selected function are subsumed under the principle that good parts have robust explanatory relations, whether those are the result of evolutionary lineages, or developmental lineages or causal-mechanical explanations. Often connectedness emerges from the fact that a part plays a role in multiple causal mechanisms (e.g., ATP synthases, discussed in the previous chapter, play roles in multiple cellular mechanism across species and domains). Additionally, parts typically have robust explanatory relations due to their lineage. Below, I explore how lineages are a source of robust explanatory relations that can help identify good parts and distinguish them from gerrymandered ones.

2.4.1 Lineage Explanations and Good Parts

Calcott (2009) outlines an influential account of a distinctive type of explanation in biology that has to do with how organisms, and their specific parts, having a common descent came to differ. Calcott (2009) terms this type of explanation “lineage explanations”. Unlike selection-based explanations that account for how particular adapted variants come to be fixed or disappear from a population, lineage explanations are concerned with how “particular modifications are responsible for the differences between organisms related by descent” (Calcott 2009, 52). Lineage explanations address how particular mechanisms work, and change, over time.

Calcott (2009) gives as examples the evolution of eyes and feathers. “Eyes” evolved from light sensitive spots to the complex organs found in numerous species through a series of discrete morphological changes within a lineage. In his second case, feather evolution is indexed to the discrete changes undergone by follicles across a lineage. Changes in the form of the feather are the result of changes in the structure of the follicles that produce it. Just as squeezing toothpaste through differently shaped nozzles creates differently shaped toothpaste, modification of follicle structure across a lineage results in different types of corresponding feathers emerging (Calcott 2009, 60).

As the above discussion shows, lineage explanations enable us to identify and delineate parts both within and across species. For instance, follicles and feathers are good parts because they have robust explanatory relations. As Calcott (2009) notes “although the target of this explanation is the changing form of feathers, the explanation is given in terms of the follicle” (Calcott 2009, 60). As such, their explanatory connection holds across a range of interventions with different follicles generating distinct feather types across a lineage, and different properties

becoming relevant across evolutionary history as the function of feathers changes (e.g., from heat regulation to flight). My connectedness condition therefore identifies a carving principle that coheres with the practice of science.

To discuss another example from the scientific literature, in order to overcome the problems of identifying parts solely by their structure and function discussed above, Arendt et al. (2019) defend an approach where the key to “to cell type identity are evolutionary changes in the ‘core regulatory complex’ (CoRC) of transcription factors, that make emergent sister cell types distinct, enable their independent evolution and regulate cell type-specific traits [...]” (Arendt et al. 2019, 744). As in the approach I defend, their research also draws on evolutionary and developmental lineages to delineate types. Crucially, Arendt et al. (2019) distinguish between evolutionary and developmental lineages. Evolutionary lineages “reflect the diversification of sister cell types through the evolution of new genetic programmes” (Arendt et al. 2019, 751). In the eye cell types they discuss (shown in Figure 2.1) this involves the diversification of rods, cones, and bipolar cells through the kind of mechanism outlined by Calcott (2009). Developmental lineages, on the other hand, “is represented by the patterns of cell division and fate decisions that ultimately deploy a differentiated cell type at a specific place and time within the organism” (Arendt et al. 2019, 751). Developmental lineages are concerned with the gene to phenotype developmental mechanism that produces a cell (or part) type. Drawing on both developmental and evolutionary lineages enriches the explanatory relations between parts and improves our knowledge of them.

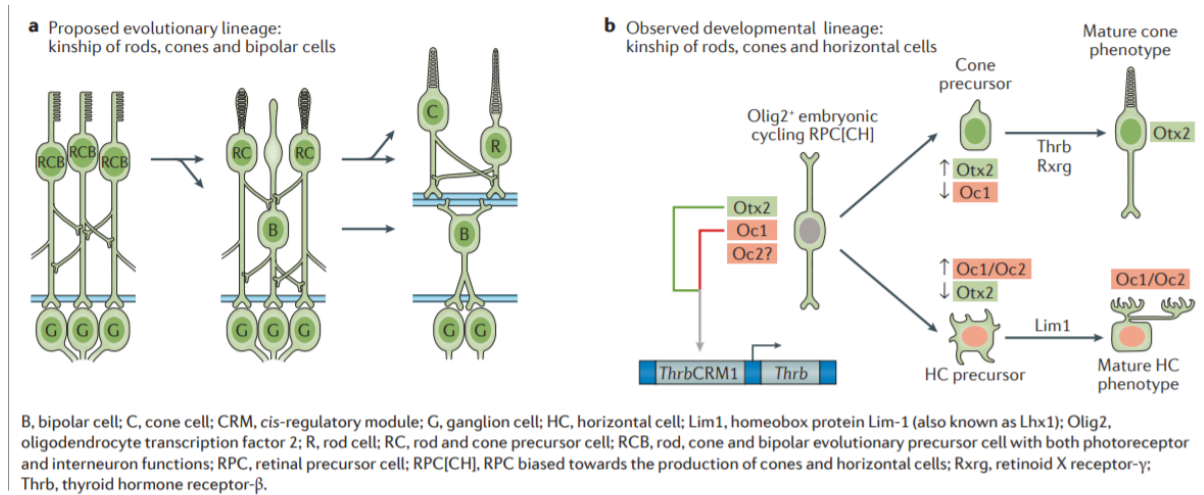


Figure 2.1: Distinction between the evolutionary and developmental lineage of cell types found in eyes [Figure from Arendt et al. 2019, 751].

My approach not only conforms to scientific practice, it also provides a principled way to distinguish between good or standard and gerrymandered parts. Below, I discuss how the connectedness condition blocks the kind of gerrymandering highlighted by Franklin-Hall (2016).

2.4.2 Connectedness and Gerrymandered Parts

Connectedness excludes gerrymandered parts such as “quarter-neurons” since these parts are explanatorily isolated. While a quarter-neuron may be specified to include all the causally efficacious properties that make it functional within a mechanism, it is explanatorily stunted. Neurons on the other hand have robust explanatory relations. Neurons figure into lineage explanations that trace the genealogy of neurons and their properties across the tree of life (Moroz 2014). They play a part in selection-based explanations that show both how neurons arose and the advantage they confer on organisms that possess them. A gerrymandered part such as “quarter-neuron,” on the other hand, is orphaned from these robust explanatory relations. Neurons fit seamlessly into the explanatory mosaic of biology in a way “quarter-neurons” do not,

making neurons and not “quarter-neurons” a good part fit for inclusion in mechanistic explanations.

The *connectedness* condition allows my account to avoid the objection raised by Franklin-Hall (2016) that the new mechanists’ appeal to scientific practice as a carving standard renders the view philosophically uninformative, a mere science reporting project. The objection arises if the store of entities is taken to be *just* what scientists have identified as parts through their methods of investigation and discovery. However, my account does not take the store of entities produced by science as philosophically primitive. Rather, I provide a justification of the store through the (partial) grounding of good parthood in the explanatory mosaic. My approach vindicates the scientific process of discovery and individuation of parts, and it does so through a philosophical analysis (i.e., the explanatory mosaic) of what makes the parts in the store good.

According to the account I defend, the causal efficacy of parts is not the sole or primary criterion by which parts are individuated. As discussed above, in the case of biology we must also consider how developmental mechanisms come to produce this part, and how this part evolved through lineages stretching into the distant past. Taken in isolation, each of these procedures seem, and indeed are to some extent, relative to the characterization of the phenomenon, which is of interest of the scientist. However, taken together they form a mutually supporting explanatory mosaic where the evolutionary and developmental lineage and causal-mechanical explanations converge on the cluster of properties that individuates a part. Parts, then, are individuated by the role they play in explanations. Parts are good when they are entrenched in the foundational theories or mechanism schemas of science. Good biological parts often have an identifiable evolutionary and developmental lineage, which may be distinct from one another. The explanatory relations between properties with a shared lineage will hold under

a wide range of interventions, i.e., is invariant, because these parts can be found in multiple species engaging in the same or different activities. It is fitting in seamlessly within this interlocking mosaic of explanations that privileges some decompositions of parts as good and preferable to alternative, non-standard decompositions.

2.5 Conclusion

The decomposition and articulation of parts is central to explanatory practice in the life sciences. This calls out for a carving principle that justifies the standard parts that are included as components of mechanistic explanations. My account of good parts provides a carving principle that secures a valuable feature of good, standard parts. It provides a principled distinction between good and gerrymandered parts that shows why the latter should not figure into our mechanism schemas. My account justifies the stable of parts special sciences use in their explanatory work by grounding good parthood in robust explanatory relations and the basic theoretical commitments of each respective science. This both respects how scientists work while defending a philosophical principle that justifies their practice. To address the insufficiency of causal efficacy as a carving principle, my account of parts states that the decomposition of a system that produces parts that do not fit into the explanatory mosaic is inferior to the one that does. And the world is such that many of the parts incorporated in mechanistic explanations fit into the explanatory mosaic. In the end, it is the richly textured explanatory mosaic produced by science that grounds what we can say about how to carve parts at nature's joints.

Chapter 3:

Mechanisms, Race, and Medicine

3.1 Introduction

Researching the sources of and potential solutions to epidemiological racial disparities¹¹ poses two challenges. On the one hand, critics of race-based studies argue that the continued incorporation of racial categories as proxies of genetic diversity reinforces the legacy of scientific racism. On the other hand, racial disparities in a number of biomedical outcomes, such as Alzheimer's disease, chronic kidney disease, low birth weights, and cardiovascular disease, among others, suggest that race remains a significant factor for understanding and potentially reversing epidemiological disparities between different racial populations in the United States, and therefore "race-based studies" are an essential component of research into these disparities (Lorusso and Bacchini 2015). Nevertheless, both advocates and critics of the race-based studies take for granted their principal mode of reasoning and investigation, namely statistical reasoning and the tools of population genetics. It is "statistical evidence of associations between variables" that is prized by the statistical reasoning approach (Matthews 2017, 1006).

The main aim of this chapter is to defend three related theses: (1) the populational approach commonly used to investigate epidemiological racial disparities (ERDs) has a number of epistemic and normative shortcomings, (2) Race (as a genetic category) does not capture the explanatorily apt populational level for epidemiology, and (3) the mechanism discovery approach is philosophically and methodologically well-suited to researching ERDs by avoiding the pitfalls of the populational approach.

¹¹ By epidemiological racial disparities I mean statistically significant differences in the incidence of disease between racialized groups. Except when discussing the views of others, I use "race" and "racialized group" interchangeably. However, see Hochman (2019) for the argument that since racialization theory is not committed to a racial ontology, "racialized groups" are conceptually distinct from "races." Since I am not committed to a racial ontology in my use of "race" (unless I am referring to another author's conception), I do not take this distinction to be a problem for my usage.

In his influential contribution to philosophy of biology, Tabery (2014) drew a distinction between *variation-partitioning* and *mechanism-elucidation* approaches to studying the relationship between genetics and human behavior. Variation-partitioning approaches seek to explain variation in a population by identifying causes of variation and *how much* variation each cause contributes. Their methodology is statistical. Mechanism-elucidation approaches, on the other hand, seek to explain *how* a given developmental process gives rise to a phenomenon by elucidating a causal mechanism. Their methodology is interventionist (Tabery 2014, 37). I defend a mechanism discovery approach that applies the distinct mechanistic reasoning and explanatory strategy to cases of epidemiological racial disparities (thesis 3). The versatility of the new mechanist approach makes it an attractive candidate for investigating ERDs where there is often an interaction of several factors from the genetic to the social.

To defend the first thesis (1), I outline the three main varieties of explanations of ERDs, which I call the racism-based explanations, genetics-based explanations, and embodiment-based explanations, and illustrate each with an example case. I argue that the dominant approach of race-based studies into ERDs, which often fall under genetics-based explanations, violate two explanatory constraints highlighted by what I call the granularity and reification problems. As I show, the granularity and reification problems pose an explanatory challenge to the prominent methodology of race-based studies of epidemiological racial disparities. This challenge stems from an inherent limitation of the populational/genetics approach in determining which variety of explanation, and subsequently what type of mechanism, is adequate.

This chapter does not defend a metaphysical view of race. For the purposes of this chapter, I assume that Black or African American, White, Asian, etc. are racialized groups and my use of “race” refers to these racialized groups and their members. Racialization, which differs

from one society to another, is a historical process of assigning individuals to different groups based on real or imagined phenotypic traits and differentially treating them in legal, political, economic, and medical spheres. This characterization might not be satisfactory for the biological race realist (see section 2), who holds that racial groups are continental populations and a proxy for human genetic diversity. But one aim of this chapter is to show that research into social components of disease mechanisms shows the salience of racism, and its contingent history, to the development of disparities in multiple epidemiological outcomes.

Furthermore, I do not claim that racial classification is never useful in “researching, diagnosing, or treating genetic disorders.” Rather, I highlight the normative and epistemic pitfalls of a primarily populational/genetics approach to investigating epistemological racial disparities. For instance, Spencer (2018) defends the usefulness of races understood as continental human populations in identifying “medically useful genetic differentiation” (Spencer 2018, 1034). He nonetheless acknowledges that “it really is a dilemma whether we should use any racial classification in a genetic way in medicine” (1034). Among some of the epistemic pitfalls is it may lead researchers to overlook social factors that have a better explanatory fit. A normative pitfall is that, as research has shown, reading about genetic diseases using racial categories raises the probability that one develops essentialist racial views, which typically leads to developing racist attitudes (Spencer 2018, 1034). What the populational/statistical approach possibly furnishes is evidence of the existence of a mechanism (Darden et al. 2018a).

In contrast to the populational approach, I argue that a mechanistic approach best satisfies the normative and epistemic constraints in investigating and potentially reversing epidemiological racial disparities. What the new mechanistic philosophy of science offers is a set of in-built epistemic norms that “guide and constrain the search for a mechanism's salient

features” (Darden et al., 2018b). The mechanistic constraints are much more robust than the populational approach, which often involves identifying genetic variants associated with a disease phenotype while black boxing the productive continuity between the two. This is because a mechanistic approach prioritizes discovery of causal mechanisms (composed of **good parts (entities)** and *activities as productive difference-makers*) over causes of variation.

The approach I defend integrates multidisciplinary experimental research involving mechanism components of different sizes (e.g., molecular, cellular, physiological, etc.). These components are nevertheless part of a single mechanism level producing a phenomenon. Following Baetu (2019), I take mechanism schemas or models that integrate experimental findings from different scientific fields to be *level neutral*. That is, as an empirical matter, the standards of evidence about causal relevance are such that it is possible to consider these findings commensurate for the purposes of schema building (Baetu 2019). The “uniformity of criteria for evaluating empirical claims about correlates, causes and mechanisms” grounds mechanism schema building drawing on findings across the natural and social sciences (Baetu 2019, 3247). Of course, there are deep philosophical questions about levels of explanation, higher level causation, constitutive levels, among others (Craver and Bechtel 2007; Potochnik 2010). However, as long as there is no empirical barrier to mechanism schema building with components drawn from different scientific fields and different sizes, integrated mechanistic research can proceed. In this chapter, I set aside important questions of inter-level causation, reductionism, and other related philosophical puzzles. What matters for our purposes, and is explored in this chapter, is the ability of mechanism components — activities and good parts— to be organized such that they contain all and only the causal factors (i.e., the causal flow from start or set up to finish or termination conditions) that produce the explanandum-phenomenon.

When it comes to the question of (biogenetic) race and biomedicine, I argue race so conceived is often not part of the causal flow in mechanisms of disease (thesis 2).

The chapter goes as follows. First, I outline biological racial realism and the disputed role of race in biomedicine and highlight its explanatory shortcomings (section 3.2). Mainly, I argue that race as the biological racial realist conceives it is not a *good part* and that the underlying genetic basis of race, whatever it may be, is often not part of the *productive difference-making* of the pathogenesis of ERDs. Second, I argue that the dominant alternative to a mechanist approach to explanation of ERDs, the populational/genetics approach, suffers from the reification and granularity problems (section 3.3). Third, I apply the new mechanist approach to my main case study: racial disparities on birth weight and highlight the virtues of this approach (section 3.4). I conclude (section 3.5) with a discussion of the diverse applications of mechanism discovery in philosophy and medicine.

3.2 The Biomedical Race Debate

The biomedical race debate concerns what role, if any, race plays in researching, diagnosing, and treating complex trait disorders. The controversy arises when race is taken to play the role of proxy for underlying genetic diversity between groups that is explanatory of phenotypic (epidemiological) diversity. Given the nature of the claim, the biomedical race debate is frequently reduced to the question of whether race is a meaningful biological category. There are broadly two families of views on the question: biological racial realism and biological racial anti-realism.

Biological racial realism (henceforth realism) holds that race is a biologically real category (Sesardic 2010; Spencer 2015, 2018; Hardimon 2017). According to realists, (1) there is

a structure to human genetic diversity that divides humans into (roughly) continental populations and (2) these continental populations (roughly) correspond to racial categories. Our racial categories therefore, according to the realist, identify an underlying biological reality. Biological racial anti-realism (henceforth anti-realism) holds that race is not a biologically real category (Kaplan and Winthers 2014; Yudell et al. 2016). Anti-realism is consistent with holding that race is a political or social entity; anti-realists merely deny that race is a meaningful or significant biological category.

The debate between realists and anti-realists hinges on whether human genetic diversity tracks divisions along continental lines that can reasonably be taken as racial; that is, whether human continental populations are identifiable clusters of genetic similarity. Rosenberg et al.'s (2002) landmark study investigating human population structure is taken by biological racial realists as the best evidence for their view. That study used a model-based clustering algorithm called *STRUCTURE* to identify major “genetic clusters” (K) based on genomic similarity.

Three things are important to note in their findings. First, in confirmation of earlier cluster studies, almost all human genetic diversity is between individuals. Rosenberg et al. (2002) find that 93-95% of genetic variation is between individuals within any cluster group. Second, the proportion of genetic diversity between a member of one cluster group and a member of another is only slightly greater than the genetic diversity between two distantly related members of the same cluster group. Finally, there are only rare cases of alleles being region-specific as “this overall similarity of human populations is also evident in the geographically widespread nature of most alleles ... region-specific alleles were usually rare, with a median relative frequency of 1.0% in their region of occurrence” (Rosenberg et al. 2002, 2381-2382). Nevertheless, *STRUCTURE* was able to identify genetic clusters that correspond to

continental groups. In particular, at $K=5$, *STRUCTURE* identified clusters correspond largely to major continental geographic regions (Africa, America, Eurasia, East Asia, Oceania). Spencer (2018) argues these clusters are identical to the Office of Management and Budget's (OMB) racial classification.

However, in order to be explanatorily apt, genetic variation at the level of racial category must correspond to variation in rates of disease incidence characterized at that level. Let us call this the *proportionality constraint* on racial explanation. The racial categories must secure an explanatory value that is missing either at higher (species) or lower (local population) organizational levels. Genetic studies must therefore demonstrate a causal-explanatory relationship between “medically relevant” genes that vary racially and disease-phenomena. I argue that they fail to do so.

3.2.1 The Realists' Race as a Gerrymandered Part

The race concepts contemporary realists such as Hardimon (2017) and Spencer (2018) defend differ significantly from historical conceptions of race. Contemporary biological racial realists conceive of race as having epidemiological or evolutionary explanatory potential. In the era of “scientific racism”, however, race was not only a potentially medical or evolutionary category. Rather, it was proposed as explanatory of the psychological, cultural, political, aesthetic, and other traits represented as belonging to racial groups.

Carl Linnaeus, in his *Systema Naturae* (1735), provided one of the first systematic taxonomies of human variation, classifying humans into *Europaeus albescens*, *Americanus rubescens*, *Asiaticus fuscus*, *Africanus nigrificulus*. Linnaeus took these to be varieties of humans (“homo variat”). The association between human variation and skin color implied by his naming

convention— yellowish Asians, blackish Africans, etc.— remained a prominent element of subsequent racial classification. Beyond physical characteristics, Linnaeus attributed traits ranging from aesthetic judgment to preference for forms of government to racial differences (Müller-Wille 2014). His characterization of human variation deeply influenced subsequent developments in anthropology and human biology (Marks 2007).

Race in the era of scientific racism was therefore part of the explanatory^x mosaic of the life sciences. Race as a biological category was taken to be robustly related to a host of human traits. I say explanatory^x, instead of explanatory, because the putative relations proposed by anthropologists and biologists in the era of scientific racism did not in fact hold. A person's race, biologically conceived, does not make a difference to whether they prefer republics to monarchies or whether they are moved by Bach's unaccompanied cello suites. The proposed relationship between race and moral, behavioral, aesthetic and other traits of social significance is now widely rejected within human and genetic sciences (Yudell et al. 2016).

As such, contemporary realists have, as one proponent calls it, a *deflationary* realism about race (Hardimon 2017). The view is deflationary because it restricts the explanatory aegis of race to phenomena related to the rise of certain populations (e.g., Native Americans from Northeastern Asians) and, perhaps, epidemiological phenomena. However, on my account, what that deflation ends up creating is an explanatorily stunted entity without the grounding in robust explanatory relations that may justify its good parthood. As I show throughout this chapter, local populations and ancestry groups are frequently more explanatorily robust than the coarser grained racial categories (Yudell et al. 2016).

I argue that race as conceived by the realist is a gerrymandered part. On Spencer's account, what makes race a biological population is its role in explaining population genetics facts such as the ones generated by cluster-based programs like *STRUCTURE*. But on Spencer's account what race explains is other facts about race. For instance, Spencer (2018) argues that the genetic similarity between Native Americans and Northeast Asians is explained by genetic drift and founder effects (Spencer 2018, 1026). My objection is not that this move is circular. Rather, I argue that the limited number of explanatory relations race (as conceived by the realist) renders it a gerrymandered part. As I argue in chapter two, what makes parts good is their membership within the explanatory mosaic of a science. And race is not part of the explanatory mosaic of biology. This is so because race, as identical to human continental populations, figures into so few of our biological explanations. Rather, it is lower-level populations groups, such as local ancestry groups, that are explanatorily fecund across a range of biological mechanism schemas. To see why, let us turn to the principal investigative tool and body of evidence realists draw on to support their claims, namely, GWAS studies.

Genome-Wide Association Studies (GWAS) are a set of tools drawn from population genetics used to identify statistical relationships between single-nucleotide polymorphisms (SNPs) and phenotypic traits. A SNP is a nucleotide, at a particular locus on a chromosome, that varies across individuals. GWAS studies scan the genome of suitably defined groups to identify whether there is SNP variation (e.g., a group of individuals with Multiple Sclerosis and a control group of individuals without the disease). GWAS studies are at the forefront of research into the relationship between genotypes and diseases. The GWAS Catalog includes GWAS studies that have found over 50,000 significant associations between genetic variants and diseases and traits (Tam 2019). Nevertheless, GWAS studies do not indicate whether these associations are causal

or spurious, or whether the implicated SNPs are actually involved in the pathogenesis of the diseases with which they are associated.

GWAS studies on epidemiological racial disparities attempt to identify SNPs associated with a disease phenotype within human sub-populations. However, GWAS studies on genetic variation and epidemiological racial disparities have turned up generally disappointing results (Kaufman et al. 2015). There are two limitations of GWAS studies into ERDs:

(1) GWAS studies into some prominent ERDs such as cardiovascular disease have found no association between genetic variation among racial groups and disease risk (Dikilitas et al. 2020);

(2) GWAS studies are constrained by a population stratification (PS) problem (Hellwege et al. 2017). PS is a significant limitation on the robustness of GWAS' association identification. PS is a result of non-random mating "due to geographic isolation of subpopulations with low rates of migration and gene flow over the course of several generations" (Hellwege et al. 2017, 2). This isolation is not caused solely by continental barriers but also as a result of linguistic, cultural, or other barriers to free and random mating. In the absence of random distribution of alleles, GWAS studies need to account for this underlying stratification in order to avoid spurious correlations.

(1) and (2) indicate GWAS studies face serious challenges to their explanatory adequacy. (2) especially drives at the limited role continental races, as opposed to more local sub-populations, play in capturing genetic variation among groups. As Hellwege and colleagues (2017) show, "the differentiation among subpopulations is detectable even when the regional differences are subtle, as has been described in Chinese and Japanese and European populations

[...] Cultural differences among populations also create stratification, even when populations inhabit the same geographical region” (Hellwege et al. 2017, 3). Of course, arguing that race is a meaningful biological population does not preclude the reality of sub-racial populations and their explanatory role. However, it is not explanatorily apt to attribute explanatory relations at lower population level to the higher-level racial population. Race (as the realist conceives it) therefore figures into far fewer explanatory relations than the more fine-grained local populations of which it is composed.

3.2.2 Race and Productive Difference-making

Race, as a biological population, is not part of the of the productive continuity of the most pressing epidemiological racial disparities. That is, race does not engage in activities, productive difference-makers¹², that bring about ERDs. In this section, I show how either social factors such as racism, or biological factors related to lower-level populations as opposed to race are explanatorily apt. This is not to deny that population-level genetic differences can play an explanatory role in medicine. Rather, I argue the population in question is only rarely racial. Given the factors that produce genetic diversity between continental populations— reproductive isolation, selection, genetic drift— also operate at much finer grain, it would be a massive coincidence if it were racial difference, as opposed to populational differences at a different grain, that accounted for a large share of epidemiological difference.

For instance, consider a case that is frequently cited by realists, namely, sickle-cell anemia. The sickle cell disease mechanism is well-characterized. Although sickle-cell anemia produced by the sickle hemoglobin allele (HbS) is widely identified as a disease affecting

¹² See chapter one for details on the Hybrid Account of Activities.

“African Americans”, there is considerable heterogeneity in its incidence within the putative African race. As Allison (2002) shows, “among tribes living close to the coast of Kenya or to Lake Victoria, the frequencies exceeded 20%, whereas among several tribes living in the Kenyan highlands or in arid country, the frequencies were less than 1%. These differences cut across linguistic and cultural boundaries and were independent of blood group markers that we documented” (Allison 2002, 2080). The source of this heterogeneity is whether malaria is prevalent in a given region. While being homozygous for the HbS allele that produces sickle-celled hemoglobin leads to debilitating disease, being heterozygous for the allele provides protection against malaria. There is therefore strong selection pressure for the allele in malaria prone regions where malaria has long been endemic (Allison 2002; Peil et al. 2010).

Piel and colleagues (2010) found “the gradual increase in HbS allele frequencies from epidemic areas to holoendemic areas in Africa is consistent with the hypothesis that malaria protection by HbS involves the enhancement of not only innate but also acquired immunity” (Piel et al. 2010, 3). Given that the African ancestors of African Americans were predominantly drawn from malaria prone regions, they unsurprisingly have higher rates of sickle-cell disease and other “genetic variants that confer resistance to malaria are associated with RBC [red blood cell] traits in African-Americans” (Ding et al. 2003, 1061). Geography and local ancestry better account for the variation in sickle-cell disease than the realist’s racial categories¹³. Therefore, the evolutionary lineages of these *localized* groups, and the distinct selection pressure that led to the prevalence of alleles that are adaptive in those specific bio-geographic environments, ground the explanatory relation between HbS frequency and continental sub-populations, not races as

¹³ Interestingly, the association between HbS frequency and malaria endemicity holds for African and Mediterranean populations but is weak in Asia and South America. Piel and colleagues (2010) hypothesize this may be due to the relative recency of the arrival of malaria to these regions, among other factors.

conceived by the realist. In section 3.4, I elucidate another case where ancestry and geographical origin cut against the racial realist approach. Before that, in the following section, I discuss the non-mechanistic populational approach to investigating epidemiological racial disparities, and its epistemic and normative shortcomings.

3.3 Race and Biomedicine: Two Approaches

In the United States, there are a number of significant epidemiological racial disparities. African Americans are two to three times more likely than Whites to develop chronic kidney disease (CKD) (Tarver-Carr et al. 2002), are twice as likely as Whites to develop Alzheimer's disease (Alzheimer's Association 2019, 333), and have higher rates of mortality from heart disease, strokes, and breast cancer, among other complex trait diseases (Goosby and Heidbrink 2013). These disparities are the subject of growing research at the intersection of race and medicine. Perhaps the most famous recent example of this is the development of BiDil¹⁴, a drug approved by the US Food and Drug Administration (FDA) to treat heart failure in African American patients (Temple and Stockbridge 2007).

There are three broad explanatory categories of research into ERDs: genetics-based explanations, racism-based explanations, and embodiment-based explanations. Genetics-based explanations take the principal causal factor in the development of ERDs to be genetic differences between racial populations. A candidate case is research in racial disparities in Alzheimer's disease (AD) discussed below. Racism-based explanations draw from the burgeoning work on social determinants of health to explain ERDs in terms of the differential

¹⁴ BiDil was approved for use in "self-identified Black patients" after a 2001 African-American Heart Failure Trial (A-HeFT) showed significant improvement in outcomes for patients on that drug cocktail (Puckrein 2006). However, the trial has been criticized for enrolling only African American subjects and for a lack of a physiological basis (a how-plausibly mechanism) for difference in drug response between races (Kahn 2005).

exposure of racial minorities to harmful social and environmental factors, such as discrimination, stress, poverty, inadequate housing, and so on. I discuss the case of ERDs in asthma to illustrate this approach. Embodiment-based explanations show how socially determined health outcomes can become embodied in complex biological, though not genetic, mechanisms. Unlike in cases where racism-based models are explanatorily adequate, these embodied disease mechanisms perpetuate ERDs even in situations where proxies for racism—discrimination, poverty, inequality of health access, etc.—have been controlled for (i.e., are not a difference-maker between the target racialized groups under study). Embodiment-based explanations appeal to cases where racism’s effects come to be embodied in racialized groups and individuals (Gravlee 2009; Sullivan 2013; Kaplan 2014). As an illustrative example of embodiment, I discuss disparities in incidence of low birth weights between White and African Americans. These explanatory strategies are not necessarily exclusive, and each has a place in research into ERDs. However, I argue the choice of which explanation is successful is best guided by a mechanism discovery approach.

Lorusso and Bacchini (2015) examine two prominent approaches taken in epidemiology and biomedical research into these disparities, which they call “race-based” and “race-neutral studies” (Lorusso and Bacchini 2015, 56). Race-based studies consider race a relevant variable in the etiology of complex diseases. The most common way race is taken up as a variable in the study design of race-based studies is as a proxy for a causally relevant factor in the production of the disease phenomenon. Usually, this causal factor is taken by many race-based studies to be *genetic*, and race therefore plays the role of proxy for genetic diversity. Lorusso and Bacchini (2015) highlight this “genetic hypothesis”, which holds that “differences in the risk of complex diseases among racial groups are largely due to genetic differences covarying with genetic

ancestry which self-identified races are supposed to be good proxies for” (Lorusso and Bacchini 2015, 57). Lewontin (1972) sparked an ongoing debate on whether races conceived as continental populations capture the structure of human genetic diversity with philosophers weighing in for (Sesardic 2010; Spencer 2015) and against (Andreasen 2004; Kaplan and Winther 2013) the genetic hypothesis. Lorusso and Bacchini (2015) argue that using self-identified races as a proxy for genetic diversity is both scientifically suspect and frequently obscures the real role played by racialization in fostering epidemiological racial disparities, namely the role of racism in differential exposure of different self-identified races to environmental and social determinants of health such as pollution, poverty, lack of education, and poor health care.

As I show below, the concerns about race-based studies are grounded in reasonable worries about reifying race as a genetic concept. The focus of research resources on finding genetic differences to explain ERDs may also engender neglect of factors that actually make a difference to ERDs. Genetics-based explanations of disease risk are nonetheless sometimes fruitfully expanded to incorporate race (as a proxy of genetic diversity) as a potential risk factor. An illustrative case of this is Alzheimer’s disease (AD). There is a robust association between the *APOE4* variant of the *APOE* gene and increased risk for cognitive decline. Apolipoprotein E4 (apoE4) is present in more than half of AD patients, making *APOE4* “the most prevalent genetic risk factor for AD” (Safieh et al. 2019). *APOE4* is one of three isomorphs of the *APOE* gene, which carries instructions for the synthesis of apolipoprotein E (apoE), an important protein in lipid metabolism and transport. The apoE4 protein, unlike the more common apoE3 variant, is less effective at transporting lipids, potentially contributing to pathological buildup of cholesterol in the brain (Safieh et al. 2019).

The initial association between *APOE4* and AD was established within White populations. But the higher incidence of *APOE4* among African Americans suggested a genetics-based explanation of the significantly higher rates of AD among African Americans (Barnes and Bennett 2015). Race-based studies into ERDs in the rate of AD have also uncovered a strong association between AD and other variants of genes involved in lipid metabolism, such as the ATP-binding cassette transporter (*ABCA7*), that are expressed in higher rates among African Americans (Reitz et al. 2013; Logue et al. 2014, Berg et al. 2019).

The upshot of the preceding discussion is that the populational/genetics approach that predominates research into AD disparities between racialized groups does not elucidate all the causally productive difference-makers and how they are organized to produce the disease phenomenon. What the populational/genetic approach provides is evidence for association between genetic factors and disease. Even when there is a well-established association between particular gene variants, such as *APOE4*, and AD risk, the relationship between genotype and phenotype in the case of complex diseases is not one of simple determination. There are a number of other genetic, comorbid, and environmental factors that determine whether possession of a gene variant associated with disease risk does in fact produce the disease phenomenon (Kaup et al. 2015). In the case of AD, education, socioeconomic status, comorbidities such as diabetes, among other factors, contribute to resilience to development of AD in individuals with high risk gene variants (Stepler and Robinson 2019). There is evidence that suggests prevalence of comorbidities is “a larger contributing factor than genetics” to AD (Stepler and Robinson 2019, 2). Evidence suggests African Americans’ lower educational attainment, higher levels of cumulative stress, and lower socioeconomic status when compared to White Americans plays a significant role in AD development (Hendrie et al. 2018). Given the large number of modifiable

risk factors that are associated with AD along with particular genetic variants, genetics-based explanations will likely fail to sufficiently account for epidemiological racial disparities in AD.

In summary, race-based studies and genetics-based explanations of ERDs reify the genetic race concept by implicitly or explicitly relying on race as a proxy for genetic diversity. Genetics-based explanations proceed by attempting to identify associations between a gene variant that predominates within a racially-defined population and disease risk. However, even with excellent candidate cases of “genetic” diseases such as AD, there are philosophical and methodological problems with how race-based studies make use of the race concept. To see why, I next outline two conceptual and scientific challenges, which I call the *granularity* and the *reification* problems, that highlight the pitfalls of a genetics-based approach that need to be avoided when engaging in race-based studies.

3.3.1 The Granularity Problem

The granularity problem faced by race-based studies reflects the difficulty in identifying just how to racially define the populations that are the relevant target for investigation and possible biomedical intervention, in the design of scientific studies and treatment protocols. As Hochman (2013) argues, following Kitcher’s (2007) initial characterization of this problem, the “grain-of-resolution problem” arises because “the appropriate grain of analysis is unclear” when it comes to the number and membership of putative races (Kitcher 2007; Hochman 2013, 345). For instance, African Americans, Black Caribbeans, and West Africans share recent common ancestry. The US Office for Management and Budget’s (OMB) scheme of racial classification, which is used across federal government agencies in the US (including the US census) and a large share of epidemiological studies to assess race membership, incorporates these populations

as a single racial group (Black/African American) (Green et al. 2002; Spencer 2018). However, the increased mortality and morbidity of African Americans in relation to White Americans and Europeans frequently does not obtain in the case of West Africans and Black Caribbeans. West Africans and recent Black immigrants to the United States do not have the same higher disease risk for hypertension, low-birth weight and premature deliveries, and Alzheimer's disease, to take three examples (Kuzawa and Sweet 2009; Valles 2012; Prince et al. 2013). Indeed, it is necessary to distinguish between White Americans and European populations (e.g., Finns) when researching diseases such as cystic fibrosis. While White Americans have a higher disease risk for cystic fibrosis compared to other racial populations in the US, Finns and Finnish migrants to the US do not (Valles 2012). The granularity problem highlights in the context of epidemiology the genetic heterogeneity within putative racial categories, raising a grain-of-analysis issue for those seeking to use "race" as a proxy for genetic diversity.

Frohlich and Potvin (2008) highlight this heterogeneity in disease risk in subpopulations in their critique of the influential "population approach" of public health research and intervention, which emphasizes maximizing harm reduction by targeting small improvements in large populations over large improvements in smaller, high-risk populations (Frohlich and Potvin 2008). The population strategy is frequently defended on grounds that it is cost-effective, as it has a single target population, and will therefore maximize the potential benefits achieved given the risks. But as Valles (2012) points out, this lumping of low-risk and high-risk populations engenders waste of resources on low-risk populations for particular diseases as a result of commitment to a dubious racial category (Valles 2012, 406).

The population approach also suffers philosophically in terms of the explanatory value of race-based research. As Root (2003) notes, race-based research in cases where there is

heterogeneity of disease risk in the population substructure is explanatorily worse than alternative approaches (discussed below) since it only applies to a subset of the population (i.e., race) that these studies take to be a proxy for a causal factor in the etiology of the disease phenomenon. The more encompassing the population is (lumping), the more it ignores the heterogeneity in disease incidence among different subpopulations racialized as Black. The less encompassing it is (for instance by stratifying the target population by ethnicity, i.e., splitting), the less race can be plausibly considered a good proxy for causal factors (such as population-specific genetic variation).

The granularity problem arises in the Alzheimer's disease case discussed above (Reitz et al. 2013; Logue et al. 2014). GWAS studies do not distinguish African Americans from African and Black Caribbean immigrants. However, this reliance on the "genetic hypothesis" leads to a potentially serious oversight. Namely, it ignores how different populations that are racialized as Black do not in fact have the same risk for developing Alzheimer's disease. Several studies of African populations have found that Africans are the least likely "continental" population-group to develop late onset Alzheimer's disease (Prince et al. 2013). Hendrie et al. (2001; 2018) found that Yoruba¹⁵ communities in Nigeria had significantly lower rates of Alzheimer's than African American communities in Indiana (Hendrie et al. 2001; 2018). The heterogeneity in Alzheimer's disease risk cuts against the strategy of finding genetic variation between White and African Americans as a causal-explanatory factor in the difference in disease risk (that is, more fine-grained population groups are epidemiologically appropriate in the case of AD). The *grain* at which racialized (sub)populations are characterized can therefore significantly alter what

¹⁵ African Americans are fairly admixed with only modest variation in ancestry and are primarily drawn from populations originating in West and West-Central Africa, including the Yoruba in Nigeria (Zakharia et al. 2009).

conclusions we can draw from race-based studies. The granularity problem highlights this explanatory tradeoff between narrow explanations that are no longer justifiably drawing on racial categories and general explanations that cast too wide a net by using race concepts and draw in subpopulations that do not display the same epidemiological phenomena.

3.3.2 The Reification Problem

The reification problem with race-based studies highlights the fact that by centering *race* as a proxy for causal factors leading to the etiology of diseases, such studies neglect the explanatorily more robust role of *racism* in determination of epidemiological outcomes. Racism is increasingly seen as an exposome, which is the totality of the environmental factors, including economic, political, and social factors, to which individuals and groups are exposed. A racist exposome significantly contributes to the worse health outcomes of African Americans (and other racialized groups exposed to similar racist exposomes) (Goosby and Heidbrink 2013). In many cases of ERDs, the exposome divides populations along racial lines with far greater significance than the relatively minor between-group genetic differences.

I call this the “reification problem” because it highlights the tendency of race-based studies to *reverse* the temporal and causal relationship between race and disease risk. It is not the fact of race that constitutes the difference-making factor in many of the epidemiological cases discussed above. Rather, the social form that “race” takes, with the well-known history of racism, exploitation and discrimination, results in significant biomedical, but not genetic, differences between racialized groups (Sullivan 2013; Kaplan 2014). The case I outline below, racial disparities in birth weight, highlights the reification problem of race-based studies and argues that the approach of mechanism discovery better captures the etiology of disease phenomena.

Lorusso and Bacchini's (2015) preferred approach is race-neutral studies. This approach does not suffer from the two problems highlighted above. Race-neutral studies do not take the significant difference between racial groups that leads to a disparity in epidemiological outcomes to be population-specific genetic differences. Race-neutral studies therefore investigate any "racial" susceptibility to disease as part of the outcome of the etiological disease mechanisms rather than as part of the cause (Lorusso and Bacchini 2015). Rather, they seek to represent the "general mechanism through which racism can chronically impact individual health" (Lorusso and Bacchini 2015, 61). This approach avoids the pitfalls of race-based studies by investigating the *mechanisms* that are operating to produce these disparities and paying careful attention to the dynamic role of racism as opposed to the fixed role of race. These race-neutral studies therefore fall under the category of racism-based explanations.

An illustrative case of a racism-based explanation is the disparity in cases of asthma. African American children are twice as likely as their White peers to develop asthma (Alexander and Currie 2017, 186). Previous studies had linked the higher rates of asthma in African American children to the greater incidence of low birth weights among African Americans. Alexander and Currie (2017) find that even accounting for birth weight differences, African American children still have significantly higher rates of asthma. They study children admitted to hospitals for asthma in New Jersey. They find that the racial difference in asthma rates is explained by residential segregation. Children, both White and Black, in Black zip codes (defined as a zip code where at least half of the children residing are African American) have higher rates of asthma than children in majority white zip codes. This is due to the fact that Black zip codes are closer in proximity to sources of outdoor pollution (highways and polluting industrial plants), have homes that are on average seven years older (and higher indoor pollution due to mold and

rodents), and a higher percentage of households with income less than \$20,000 (Alexander and Currie 2017, 194). The reason African American children have higher rates of asthma is because they disproportionately live in Black zip codes where they are exposed to these environmental factors (94 out of 676 zip codes are Black in their analysis) (Alexander and Currie 2017, 195). What makes a difference to ERDs in asthma is therefore the social and historical facts of racism and residential segregation and not putative differences between races. Both White and Black children in Black zip codes had higher rates of asthma than their peers of both racial groups in non-Black zip codes. What makes Black children more likely to have asthma than White children are the facts of residential segregation and the attendant factors discussed above (Alexander and Currie 2017).

The move to incorporate race as a proxy for a causal factor in the development of ERDs is partly driven by the downside risk of ignoring race as a relevant variable when the outcomes are clearly racialized. As we have seen, critics charge that racial categories are unwarranted as proxies for genetic diversity. This is not to suggest that race-based studies are inherently wedded to the genetic hypothesis. However, race-based studies frequently make use of race as a proxy for genetic diversity. Finding a relationship between genotypes and phenotypes, in this case SNPs that predominate in a racialized group (genotype) and disease (phenotype), remains the preeminent approach in race-based studies. These suffer from the granularity and reification problems in the context of biomedicine, on top of the challenges to genetic race concepts forwarded in the scientific and philosophical literature. Racism-based explanations avoid granularity and reification problems by following a “search for mechanisms” research process. A large share of the literature on social determinants of health and the effect of racism and low socioeconomic status on health outcomes is mechanistic (Goosby et al. 2018). As Matthews

(2017) notes “... reasoning mechanistically about a phenomenon positively influences scientific hypotheses [...]” (Matthews 2017, 1003). It does so because a mechanistic approach foregrounds the importance of identifying the entities and causally productive activities driving the etiology of puzzling phenomena. This approach disciplines the process of discovery by illuminating *how* the relevant causal factors fit together structurally and temporally, not just *what* those factors may be.

3.4 Discovering Mechanisms

Racial Birth Weight Disparities in the United States

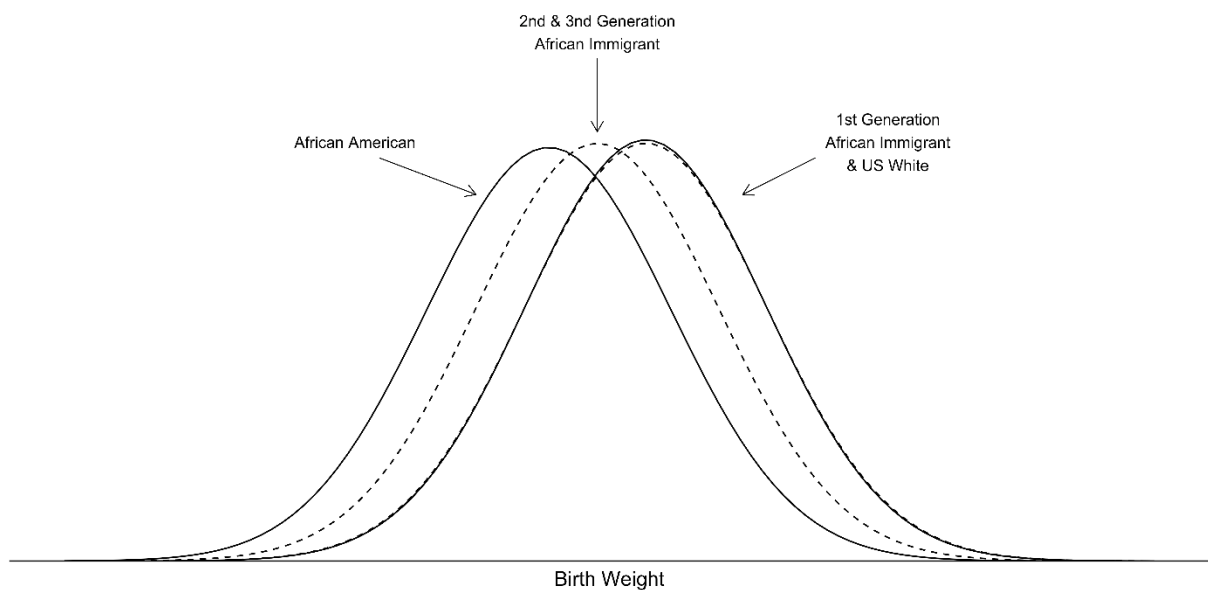


Figure 3.1: Schematic representing the comparative mean birth weights of US White, US Black, and recent African immigrants. [Note: figure from Kuzawa and Sweet 2009, 9]

Figure 3.1 represents racial disparities in birth weight (races are based on self-report). African-American women are twice as likely as their White peers to have low birth weight offspring (Kuzawa and Sweet 2009; Goosby and Heidbrink 2013). But note the even more puzzling phenomenon that first generation African immigrants have birth weights comparable to

US Whites. In order to explain the phenomenon of racial disparities in birth weight, let us build a mechanism schema from existing research into the topic.

Multiple studies have demonstrated low birth weight among African Americans. Centers for Disease Control and Prevention data shows that African American women are twice as likely to have low birth weight births as US Whites, and 2.69 times more likely to have very low birth weight (LBW) births (CDC 2005; Kuzawa and Sweet 2009). This disparity has been consistent over several decades (Kramer et al. 2006; Kuzawa and Sweet 2009). The racial birth weight disparity is made all the more puzzling by the fact that recent African immigrants born overseas have birth weight distributions that are nearly identical to US Whites (Forna et al. 2003). However, a study conducted in the State of Illinois found that over subsequent generations the descendants of recent African immigrants had birth weights that were approaching the African American mean (Figure 3.1; Collins et al. 2002; Kuzawa and Sweet 2009).

In accounting for the racial disparities in birth weight in the US, a primarily *genetic* mechanism seems highly implausible. The birth weight of first-generation African migrants is not significantly different from White American birth weights. Surprisingly, second and third generation African immigrants have birth weights whose mean begins to approach the birth weight of African Americans and after several generations there is no statistically significant difference in their mean birth weights (Kuzawa and Sweet 2009, 5). Furthermore, several studies have shown the low heritability of birth weight (Vlietinck et al. 1989; Whitfield et al. 2001). The change in the African migrant birth weight figures across generations suggests that there are environmental factors within the United States that affect both groups (who are racialized as Black regardless of national origin).

That environmental factor is the different *exposome* to which White and African Americans are subject. In order to adequately explain low birth weight in African Americans we must include the social level and the role of structural racial inequality. The link between racial discrimination, resulting chronic stress, and low birth weight/preterm birth is well established (Goosby and Heidbrink 2013, 636). Chronic stress also leads to low birth weight among White women. However, when the stressor, such as poverty or low income, is removed, White birth weights return to normal levels. What is surprising is that equivalent rises in income for African American women did not have a statistically significant effect on birth weight, i.e., the birth weights remained as low as before (Goosby and Heidbrink 2013, 638). It is the fact that in such cases proxies for racism-- discrimination, unemployment, low socio-economic status, etc.—are not explanatory that justifies looking beyond simple racism-based explanations to the disease mechanisms of embodiment-based explanations. Below, I discuss the evidence for embodiment of disparities in birth weight between groups racialized as Black and White in the United States.

Kuzawa and Sweet (2009) argue that chronic stress experienced by African Americans creates an intrauterine environment which subtly alters fetal development, most importantly, increased maternal cortisol, insulin, and blood pressure. The entity that engages in the difference-making activities in this process is cortisol which, upon being absorbed through the placenta by the fetus, acts to induce a slow-down in fetal growth and in high enough doses to induce pre-term birth (Kuzawa and Sweet 2009; Goosby and Heidbrink 2013).

Furthermore, there is evidence, although not conclusive for a how-actually mechanism, that some of these fetal alterations may be epigenetic. Epigenetics refers to modifications of the rate of expression of genes without any alteration to the nucleotide sequence of DNA. The most commonly studied mechanisms of epigenetic change are histone modification, which alters the

proteins around which the DNA is wrapped, and DNA methylation which adds a methyl group to CpG regions of DNA, effectively silencing them (Kuzawa and Thayer 2011). Recent studies have established a relationship between certain types of exposures and epigenetic changes, including harmful environmental exposures resulting in disease-promoting epigenetic changes (Bagby et al. 2019).

3.4.1 Assessing the Evidence

There are three crucial lines of evidence that may be brought together to propose a mechanism for low birth weight in African Americans: (i) the differential exposure of African American women to psychosocial stressors, (ii) the effect of acute stress on the metabolism and physiology of African American women, and (iii) the resulting embodiment of these cumulative stressors in African Americans. I will review the evidence for each in turn.

It is well-established that African Americans are disproportionately exposed to high degrees of social stressors (Sternthal et al. 2011). This is due in large part to racial discrimination and economic inequality (Wilson and Rodgers 2016). African American women disproportionately work in low wage jobs, are more likely to live in poverty and extreme poverty in comparison to their White counterparts and have little wealth (Wilson and Rodgers 2016). Even middle-class African Americans “are more likely to live in conditions where they are exposed to, or in close proximity, to concentrated disadvantage, high unemployment rates, pollution, violent crime, and poor housing conditions” (Goosby and Heidbrink 2013, 631-632). All of these factors contribute to the differential exposure of African American women to multiple social stressors.

The impact of stress on health is well-documented. *Psychosocial stress* is a productive difference-maker, i.e., an activity, in a number of disease mechanisms. Among the changes it produces are rapid cellular degradation (Epel et al. 2004) and elevated cortisol and blood pressure. These in turn increase the risk of developing diabetes and heart disease (McEwen and Gianaros 2010). Psychosocial stress induces excess production of the stress hormone cortisol. And there is evidence that cortisol produced due to antenatal stress is transferred through the placenta into the fetal bloodstream (Zijlmans et al. 2015; O'Donnell and Meaney 2017). Cortisol is a hormone that regulates the hypothalamic-pituitary-adrenal axis (HPA) (responsible for maintaining homeostasis) response to stress along with the maintenance of pregnancy and the onset of birth (Stewart et al. 2015). Cortisol and corticotropin-releasing hormone (CRH) rise under normal conditions during the course of a pregnancy. However, excess maternal CRH increases the risk of premature and/or low birth weight births (Phillips et al. 1998; Goosby and Heidbrink 2013) by producing poorly vascularized placentas and thereby restricting intrauterine growth (Stewart et al. 2015).

Finally, there is the transgenerational aspect of low birth weight in African Americans. This is the part of the mechanism for which there is the least evidence, and which is drawing attention in research. There are two factors that may contribute to the fact that low birth weight births are transgenerational: first, the perpetuation of the environmental factors (e.g., psychosocial stressors) which lead to low birth weight offspring and second, epigenetic modifications that increase the risk of low birth weight offspring having offspring of their own with the same low birth weight phenotype even in the absence of the initial causally productive exposome. Evidence for the first factor is discussed above. As for the second, there is evidence fetal programming is responsible for perpetuating low birth weight outcomes even in the absence

of the original inducing environment, although the evidence is limited and the exact mechanism for this is not currently known (Drake and Walker 2004; St-Pierre et al. 2012; Scholaske et al. 2018). McDade (et al. 2019) find evidence for a relationship between low socioeconomic status (SES) and epigenetic changes in a number of CpG sites across a large portion of the genome. DNA methylation was over-represented in sites associated with skeletal development, immune function, and development of the nervous system (McDade et al. 2019). Further research is needed to find out what mechanisms, if any, are mediating these effects.

3.4.2 Building the Mechanism Schema

We now have a how-plausibly mechanism for low birth weight among African Americans (Figure 3.2) [productive difference-makers (*activities*) in italics and **entities** (good parts) in bold]. Acute stress due to *racial discrimination* and structural inequality *induces* a change in the metabolism and physiology of African American **women**. This creates an intrauterine environment with *increasing* **cortisol**, **insulin**, and blood pressure (BP). Intrauterine and fetal metabolism of maternal **cortisol** and **corticotropin-releasing hormone (CRH)** *restricts* growth of the **fetus**. Ongoing stress reinforces the mechanism¹⁶. Even if the maternal exposome were to have reduced stress, fetal programming potentially makes low birth weight birth transgenerational, although more research is needed on this front (Scholaske et al. 2018).

Racial discrimination, then, is a productive difference-maker in this mechanism while the underlying genotype of women racialized as Black is not. It drives the changes in the mechanism that brings about the subsequent steps. Other activities may bring rise to the stress that leads to the physiological changes described in the mechanism schema. Nonetheless, *racial*

¹⁶ See discussion of evidence for details.

discrimination is a social activity whose occurrence has been shown to bring about the changes in maternal physiology that is a component of the etiology of LBW births among Black American women.

The components of this mechanism are drawn from different size levels, ranging from the molecular (cortisol) to the social (discrimination). However, the mechanism is level neutral. That is, the components of the mechanism are at the same mechanism level. This is due to the single productively continuous causal flow from start to the disease-phenomenon. The mechanism discovery approach does not privilege any particular component or size level, incorporating all the parts that engage in productive difference-making. This ensures no causally relevant factor is left out by focusing on a single factor, such as genetics, that may be too causally distant and restricted to adequately account for the phenomenon in question (Reimers et al. 2019). And given the level neutrality and mutual compatibility of scientific evidence drawn from different scientific fields (Baetu 2019), mechanism discovery is both a feasible and attractive explanatory approach to the low birth weight, and other, disease-phenomena.

This low birth weight mechanism is an illustrative case of embodiment-based explanations. As we have seen, although West Africans, Black Caribbeans and African Americans share genetic ancestry, these three subpopulations do not share similar health outcomes when compared to White Americans. Race (as a proxy for some genetic variable) therefore does not *productively make a difference* to the occurrence of the low birth weight disparity. Rather, by allowing the phenomenon to guide us we see that genetic race concepts are inadequate. The *exposome* to which African Americans are subject productively makes a difference to racial birth weight disparities in a way that the underlying putative genetic referents of race do not. Instead of meaningful biological differences leading to racial classification, the

social process of racial categorization ends up producing meaningful biological differences. In other words, racism acts as a start-up condition of a biological mechanism that produces meaningful biological differences between socially defined races (i.e., race comes to be “embodied”).

What makes mechanism discovery a promising approach for the explanation of and intervention on epidemiological racial disparities is that it provides productive continuity between proposed causal factors (racism, genetics, embodiment) and disease incidence by showing how the different parts of the mechanism, entities and activities, interact to produce the phenomena. As we see with the birth weight case, mechanism discovery avoids the granularity and reification problems. By identifying the exposome as the productive difference-maker, either in producing the disease or interacting with genetic factors to mediate disease risk, there is no need to lump or split different subpopulations racialized as Black in the United States in order to articulate genetically similar populations. In the low birth weight case, it shows the correct direction of the production of the low birth weight phenomenon from the social fact of racism to the biomedical racial disparities. Furthermore, mechanism discovery does not detract from, but rather enhances, our ability to propose interventions to arrest and reverse these disparities by, for instance, identifying the role of productive difference-makers such as a racist exposome and fetal programming.

3.5 Conclusion

Regardless of whether there is a scientifically useful structure to human genetic diversity that is identical or corresponds to race, that structure is not apt to explain differences in disease risk. Populational/statistical association studies lack the explanatory constraints to illuminate the

causal structure of racialized differences in disease risk. They are at best preliminary evidence of existence of etiological mechanisms of disease.

Mechanism discovery is a philosophically well-developed and fruitful approach to investigate how epidemiological racial disparities are created and maintained. The prevailing reliance on discovering statistical association between putative causal factors, genetics in particular, and disease risk has serious explanatory limitations. Using race as a proxy for genetic diversity opens up race-based studies to the granularity and reification problems. By expanding the new mechanist concept of activities to a view of activities as productive difference-makers, and developing an account of good parthood, the mechanism discovery approach I propose addresses a prominent critique of the new mechanism approach and shows the explanatory virtues of applying mechanism discovery to epidemiological racial disparities. Constructing biomedical mechanism schemas that incorporate and highlight the role of productive difference-makers enables the identification of all and only the actual difference-makers to disease incidence. The kinds of mechanisms that produce racial disparities may have genetic, biological, and/or social productive difference-makers, making mechanism discovery a potentially interdisciplinary effort in charting the mechanism schemas of racial epidemiological difference. Furthermore, mechanism discovery could be useful in designing policy and medical interventions to address racial disparities in health (Efstathiou 2012; Hardimon 2013), among other spheres, as the philosophy of mechanisms is increasingly being used in medicine (Russo and Williamson 2007; Darden et al. 2018b; Parkkinen et al. 2018; Kennedy and Malanowski 2019).

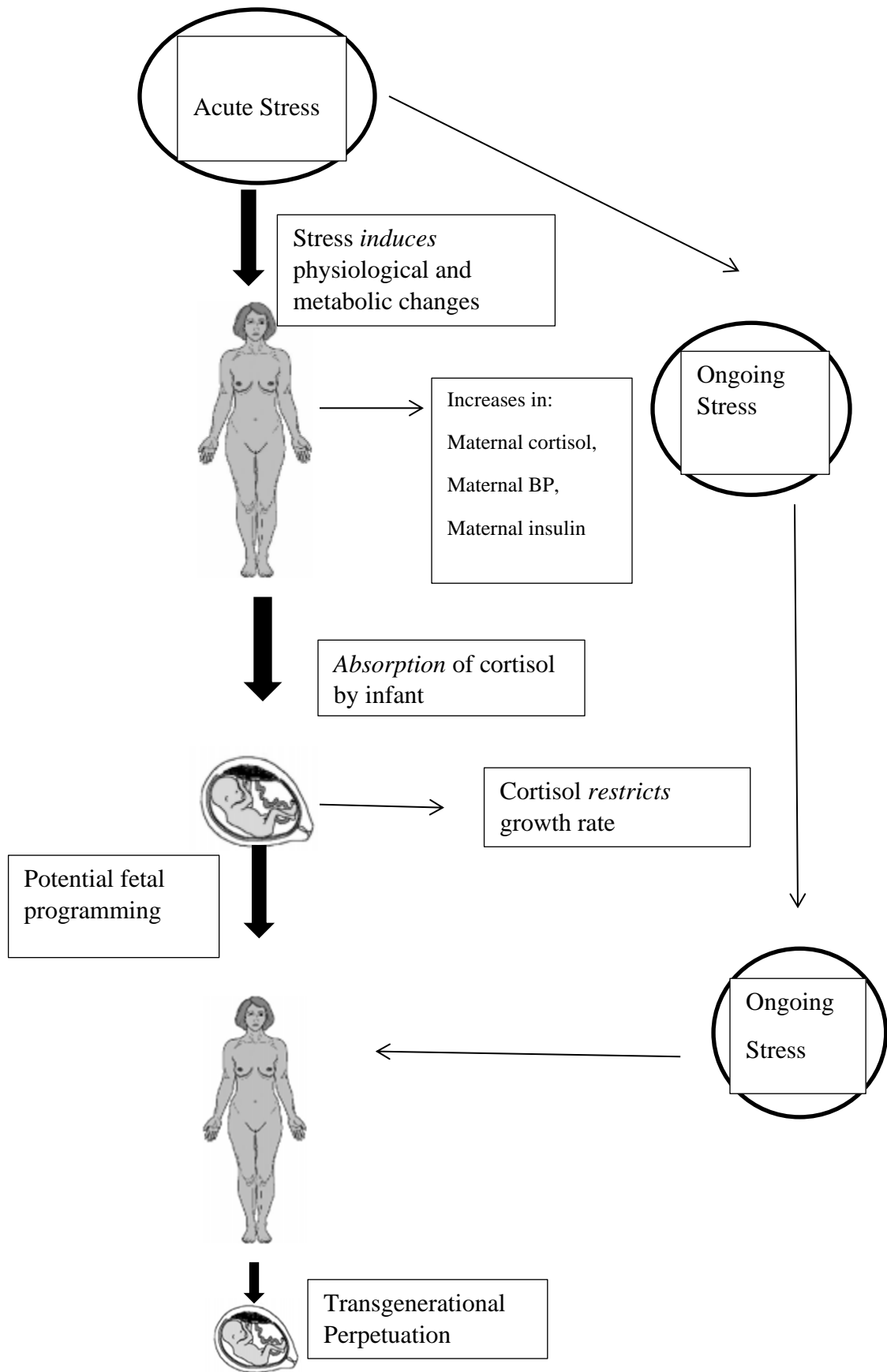


Figure 3.2: Diagram of the mechanism schema for low birth weight births in African American women, details of which are discussed above. [Note: adapted with modifications from a pathways model appearing in Goosby and Heidbrink 2013]

Conclusion

In chapter one, I developed the Hybrid Account of activities (HAA) as productive difference-makers, bringing together the production and difference-making accounts of causation. HAA allows us to distinguish the causal components of mechanisms —activities— from both irrelevant side-effects and causal influences that are merely background or enabling conditions. The account of activities as productive difference-makers has rich potential applications to outstanding questions in new mechanistic research. For instance, it can be applied to resolve the question of descriptive sufficiency. The inclusion of all and only productive difference-makers, and the parts that engage in them, can be used as a principle of when the details of a mechanism schema are enough to be adequate and complete. That is, HAA can be extended into the analysis of abstraction and idealization in science. when we abstract, we remove details from our explanations. When we idealize, we add false details to our explanations. There are deep philosophical questions about what makes abstraction and idealization valuable in science. This question is no less pressing for the new mechanistic philosophy of science, which must have an account of when the details of a mechanism schema are enough (Craver and Kaplan 2020). The union of production and difference-making in HAA provide resources for identifying which details we may remove, collapse, or idealize without rendering our mechanism schemas explanatorily inadequate.

The Hybrid Account also makes strides in the philosophy of experimentation as an explication of the discovery of new activities. For instance, in their landmark study demonstrating a new sequencing technology, Braslavsky and colleagues (2003) demonstrated a method to study the activity of a DNA polymerase (i.e., *replication*) “at the single molecule level

with single base resolution and a high degree of parallelization, thus providing the foundation for a practical single molecule sequencing technology” thereby avoiding the high cost, slower speed, and lower sensitivity of previous conventional sequencing methods (Braslavsky et al. 2003, 3960). They performed variations of experiments with and without the polymerase to establish that their methodology produces a polymerase whose activity (*replication*) produces nucleotides with “high fidelity”, fidelity being a measure of the accuracy of a polymerase. They thus tested HAA conditions (i) through (iii) to establish the viability of their method.

The method of single-molecule sequencing was in turn used by Öz and colleagues (2020) to characterize a novel activity called *DNA bridging*, which is crucial in DNA double-strand break (DSB) repair (Öz et al. 2020). They show that two protein molecules, phosphorylated CtIP and MRN complex, hold the two DNA strands together (hence “bridging”) to facilitate repair. Earlier experimental approaches had shown contradictory results about the role of CtIP in DSB repair. They, however, “used a nanofluidics-based *single-molecule approach* to characterize the *robust DNA-binding activity* of phosphorylated wtCtIP along with a library of CtIP variants, with the aim of providing insight into how CtIP can play an important role in structuring two proximal DNA ends in DSB repair” (Öz et al. 2020, 21410 [emphasis added]). The series of experiments they perform is to identify and individuate their proposed activity, and their explanatory standard, precisely captures those incorporated into HAA. HAA can therefore be fruitfully explored as an account that philosophically elucidates the discovery of activities, is consistent with scientific practice, and extends the philosophical analysis of experimental science.

In the second chapter, I defended an account of good parts that grounds why the parts that form the regular entities of science— enzymes, neurons, cells, hearts, and so on— are the well-identified and individuated components of mechanism schemas. It distinguishes between good

and gerrymandered parts, the latter of which might have activity-enabling properties but are nonetheless not part of the store of entities used by scientists. My account philosophically justifies the store of parts special sciences use in their explanatory work by grounding good parthood in robust explanatory relations that form an explanatory mosaic. The account I defend privileges the explanatory work of science in determining what the ontological structure of the world is; and it provides a philosophical account of how we have access to that structure through the building of the explanatory store and explanatory mosaic. My account added to the activity-enabling and material conditions, which are insufficient to exclude gerrymandered parts, and recruits the explanatory mosaic of science to ground good parthood.

My account of good parts intersects with the theory of natural kinds. In so far as my account of good parts identifies those parts that can be components of mechanisms, it raises the question of whether good parts are natural kinds. Given that my account does not provide necessary and sufficient conditions for good parthood, it fails to qualify for the metaphysically heavy conception of natural kindhood (such as kind essentialism). In future work, I plan to explore how good parts are *explanatory kinds* that enable not merely the prediction but also the mutually supporting explanation of their (i.e., good parts') properties.

Furthermore, my analysis of good parts can be extended to complex, aggregative parts such as populations. An account of populations as components of mechanisms has application in evolutionary biology, ecology, and social sciences, among others (Ylikoski and Kuorikoski 2010; Paslaru 2014; DesAutels 2016). Social mechanisms are a burgeoning area of research in fields ranging from epidemiology to economics (Little 2011). An account of good parts that grounds populations' suitability as components of mechanisms would advance the integration of explanations across the life and social sciences, a topic I intend to explore.

In the third chapter, I defended the mechanism discovery approach to biomedical explanation. The approach applied the accounts developed in earlier chapters (of activities and good parts) to central questions at the intersection of philosophy of race and medicine. The chapter highlighted the epistemic and normative problems arising from the use of race as a proxy for underlying human genetic diversity, and the populational/statistical approach that frequently makes such use of race. As an alternative, I proposed mechanism discovery as an explanatorily attractive approach to investigating the relationship between race and biomedicine. The approach I defended in this chapter does not reify race or focus on the wrong population grain in identifying disease-phenomena. The process of building mechanism schemas disciplines the search of causal factors by placing a high explanatory value on productive continuity. This limits the normative and epistemic risks highlighted by the granularity and reification problems. A multidisciplinary approach to schema building that seeks to represent the full causal flow that produces a disease-phenomenon best captures our core desiderata in explaining, predicting, and treating disease, in particular those displaying racial disparities.

The mechanism discovery approach has fruitful applications not just in medicine but potentially in social sciences such as economics as well. For instance, one can extend the mechanism discovery approach and argue that randomized controlled trials (RCTs) are inferior to mechanism discovery in establishing causality in economics. Developmental economists in particular have been using RCTs to investigate effective interventions. Banerjee, Duflo, and Kremer's 2019 Nobel Prize in Economics, awarded based on their work with RCTs, has brought the debate about RCTs and their role in establishing causality to the fore (Stevano 2019). Just as in the case of medicine (Russo and Williamson 2007), a mechanism discovery approach

potentially avoids the normative and explanatory pitfalls critics of RCTs have identified (e.g., Cartwright 2010).

The approaches I have developed in this dissertation, therefore, have fruitful applications to central questions at the intersection of philosophy, science, and medicine.

References

- Adeva-Andany, M. M., González-Lucán, M., Donapetry-García, C, Fernández-Fernández, C, and Ameneiros-Rodríguez, E (2016) Glycogen metabolism in humans. *BBA clinical* (5): 85–100.
- Agius, L (2015) Role of glycogen phosphorylase in liver glycogen metabolism. *Molecular Aspects of Medicine* (46): 34-45.
- Alexander D, Currie J (2017) Is it who you are or where you live? Residential segregation and racial gaps in childhood asthma. *Journal of Health Economics* (55): 186-200.
- Allison, AC (2002) The discovery of resistance to malaria of sickle-cell heterozygotes. *Biochemistry and Molecular Biology Education* 30(5): 279-287.
- Alzheimer's Association (2019) Alzheimer's disease facts and figures. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* 15(3): 321-387.
- Andreasen RO (2004) The cladistic race concept: a defense. *Biology and Philosophy* 19(3): 425-442.
- Anscombe, GEM (1981) Causality and determination, in *Metaphysics and the Philosophy of Mind, The Collected Papers of G. E. M. Anscombe*, v: 33-147. Minneapolis, MN: University of Minnesota Press.
- Arendt, D, Bertucci, PY, Achim, K, and Musser, JM (2019) Evolution of neuronal types and families. *Current Opinion in Neurobiology* 56: 144-152.
- Baetu, TM (2019) On pain experience, multidisciplinary integration and the level-laden conception of science. *Synthese* 196(8): 3231-3250.
- Bagby SP, Martin D, Chung ST, Rajapakse N (2019) From the outside in: biological mechanisms linking social and environmental exposures to chronic disease and to health disparities. *American Journal of Public Health* 109(S1): S56-S63.
- Barnes LL, Bennett DA (2015) Dementia: cognitive resilience in APOE* ε4 carriers—is race important? *Nature Reviews Neurology* 11(4):190-191.
- Bechtel, W and Richardson, RC 1993 *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton, NJ: Princeton University Press.
- Bechtel, W and Abrahamsen, A 2005 Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 36(2): 421-441.

- Berg, CN, Sinha, N, & Gluck, MA (2019) The effects of APOE and ABCA7 on cognitive function and Alzheimer's disease risk in African Americans: a focused mini review. *Frontiers in human neuroscience* 1(387):1-7.
- Berg JM, Tymoczko JL, Stryer L (2002) *Biochemistry*. 5th edition. New York: W H Freeman; Chapter 21, Glycogen Metabolism. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK21190/>
- Braslavsky, I, Hebert, B, Kartalov, E, and Quake, SR (2003) Sequence information can be obtained from single DNA molecules. *Proceedings of the National Academy of Sciences* 100(7): 3960-3964.
- Bhagal, HS (2020) Coincidences and the grain of explanation. *Philosophy and Phenomenological Research* 100(3): 677-694.
- Bogen, J (2008) Causally productive activities. *Studies in History and Philosophy of Science Part A* 39(1): 112-123.
- Bruce, HS, and Patel, NH (2020) Knockout of crustacean leg patterning genes suggests that insect wings and body walls evolved from ancient leg segments. *Nature Ecology & Evolution* 4.12: 1703-1712.
- Calcott, B (2008) Lineage explanations: explaining how biological mechanisms change. *The British Journal for the Philosophy of Science* 60(1): 51-78.
- Carlen, PL, Gurevich, N and Durand, D (1982) Ethanol in low doses augments calcium-mediated mechanisms measured intracellularly in hippocampal neurons. *Science* 215(4530): 306-309.
- Cartwright, N (2010) What are randomised controlled trials good for? *Philosophical Studies* 147:59–70. DOI 10.1007/s11098-009-9450-2
- Centers for Disease Control and Prevention (CDC) (2005) Racial/ethnic disparities in infant mortality--United States, 1995-2002. *MMWR. Morbidity and Mortality Weekly Report* 54(22): 553.
- Collins JW, Wu SY, David RJ (2002) Differing intergenerational birth weights among the descendants of US-born and foreign-born whites and African Americans in Illinois. *American Journal of Epidemiology* 155(3): 210-216.
- Craver, CF (2001) Role functions, mechanisms, and hierarchy. *Philosophy of Science* 68: 53-74.
- Craver, CF (2007) *Explaining the Brain: Mechanisms and the Mosaic unity of Neuroscience*. New York: Oxford University Press.

- Craver, CF and Bechtel, W (2007) Top-down causation without top-down causes. *Biology and Philosophy* 22 (4): 547-563.
- Craver CF, and Darden L (2013) *In search of mechanisms: discoveries across the life sciences*. University of Chicago Press, Chicago.
- Craver, CF, and Kaplan, DM (2020) Are more details better? On the norms of completeness for mechanistic explanations. *The British Journal for the Philosophy of Science* 71(1): 287-319.
- Darden, L (2002) Strategies for discovering mechanisms: schema instantiation, modular subassembly, forward/backward chaining. *Philosophy of Science* 69(Proceedings): S354-S365.
- Darden, L (2006) *Reasoning in Biological Discoveries: Essays on Mechanisms, Interfield Relations, and Anomaly Resolution*. New York: Cambridge University Press.
- Darden L, Pal LR, Kundu K, Moul J (2018a) The product guides the process: discovering disease mechanisms. In: Danks D, Ippoliti E (ed). *Building theories: heuristics and hypotheses in the sciences*. Cham, Switzerland: Springer International Publishing, pp.101-117.
- Darden L, Kundu K, Pal LR, Moul J (2018b) Harnessing formal concepts of biological mechanism to analyze human disease. *PLoS Computational Biology* 14(12), e1006540.
- Daubner S, Le T, and Wang S (2011) Tyrosine hydroxylase and regulation of dopamine synthesis. *Archives of Biochemistry and Biophysics* 508(1): 1-12.
- Deckers-Hebestreit, G, and Altendorf, K (1996) The F0F1-type ATP synthases of bacteria: structure and function of the F0 complex. *Annual review of microbiology* 50(1): 791-824.
- DesAutels, L (2016) Natural selection and mechanistic regularity. *Studies in History and Philosophy of Biological and Biomedical Sciences* 57:13-23.
- Dikilitas, O, Schaid, DJ, Kosel, ML, Carroll, RJ, Chute, CG, Denny, JA., ... & Kullo, IJ (2020) Predictive utility of polygenic risk scores for coronary heart disease in three major racial and ethnic groups. *The American Journal of Human Genetics* 106(5): 707-716.
- Ding, K., de Andrade, M., Manolio, T. A., Crawford, D. C., Rasmussen-Torvik, L. J., Ritchie, M. D., ... & Kullo, I. J. (2013). Genetic variants that confer resistance to malaria are associated with red blood cell traits in African-Americans: an electronic medical record-based genome-wide association study. *G3: Genes, Genomes, Genetics* 3(7): 1061-1068.
- Drake AJ, Walker BR (2004) The intergenerational effects of fetal programming: non-genomic mechanisms for the inheritance of low birth weight and cardiovascular risk. *Journal of Endocrinology* 180(1): 1-16.

- Efstathiou S (2012) How ordinary race concepts get to be usable in biomedical science: an account of founded race concepts. *Philosophy of Science* 79(5): 701-713.
- Epel ES, Blackburn EH, Lin J Dhabhar FS, Adler NE, Morrow JD, Cawthon RM (2004) Accelerated telomere shortening in response to life stress. *Proceedings of the National Academy of Sciences of the United States of America* 101(49): 17312-17315.
- Forna F, Jamieson DJ, Sanders D, Lindsay MK (2003) Pregnancy outcomes in foreign-born and US-born women. *International Journal of Gynecology & Obstetrics* 83(3): 257-265.
- Franklin-Hall LR (2016) New mechanistic explanation and the need for explanatory constraints. In: Aizawa K, Gillett C (ed). *Scientific composition and metaphysical ground*. Palgrave Macmillan, London. pp.41-74.
- Frohlich KL, Potvin L (2008) Transcending the known in public health practice: the inequality paradox: the population approach and vulnerable populations. *American Journal of Public Health* 98(2): 216-221.
- Galen (1968) *Galen on the Usefulness of the Parts of the Body: Peri Chreias Moriōn [Romanized Form] De Usu Partium*. Edited by Margaret Tallmadge May. Cornell Publications in the History of Science. Ithaca, N.Y.: Cornell University Press.
- Glennan, S (1996) Mechanisms and the nature of causation. *Erkenntnis* 44(1): 49-71.
- Glennan, S (2017) *The New Mechanical Philosophy*. New York: Oxford University Press.
- Godfrey-Smith, P (2009) Causal pluralism. In Helen Beebe, Peter Menzies & Christopher Hitchcock (eds.), *The Oxford Handbook of Causation*. Oxford University Press. pp. 326—337.
- Goosby BJ, Heidbrink C (2013) The transgenerational consequences of discrimination on African-American health outcomes. *Sociology Compass* 7(8): 630-643.
- Goosby BJ, Cheadle JE, Mitchell C (2018) Stress-related biosocial mechanisms of discrimination and African American health inequities. *Annual Review of Sociology* 44: 319-340.
- Gravlee CC (2009) How race becomes biology: embodiment of social inequality. *American journal of physical anthropology* 139.1: 47-57.
- Green CR, Cupples LA, Go R, Benke KS, Edeki T, Griffith PA, Williams M (2002) Risk of dementia among white and African American relatives of patients with Alzheimer disease. *JAMA* 287 (2): 329-336.
- Griffiths, PE (2006) Function, homology, and character individuation. *Philosophy of Science* 73(1): 1-25.

- Hardimon MO (2013) Race concepts in medicine. *Journal of Medicine and Philosophy* 38(1): 6-31.
- Hardimon, M. O. (2017). *Rethinking race*. Cambridge, MA: Harvard University Press.
- Hellwege, J. N., Keaton, J. M., Giri, A., Gao, X., Velez Edwards, D. R., & Edwards, T. L. (2017). Population stratification in genetic association studies. *Current protocols in human genetics*, 95(1): 1-22.
- Hempel, CG, and Oppenheim, P (1948) Studies in the logic of explanation. *Philosophy of Science* 15(2): 135-175.
- Hendrie HC, Ogunniyi A, Hall KS, Baiyewu O, Unverzagt FW, Gureje OW, Gao S (2001) Incidence of dementia and Alzheimer disease in 2 communities: Yoruba residing in Ibadan, Nigeria, and African Americans residing in Indianapolis, Indiana. *JAMA* 285(2): 739-747.
- Hendrie, H. C., Smith-Gamble, V., Lane, K. A., Purnell, C., Clark, D. O., & Gao, S. (2018). The association of early life factors and declining incidence rates of dementia in an elderly population of African Americans. *The Journals of Gerontology: Series B*, 73(suppl_1): S82-S89.
- Hochman A (2013) Against the new racial naturalism. *The Journal of Philosophy* 110(6): 331-351.
- Hochman, A (2019) Racialization: a defense of the concept. *Ethnic and racial studies* 42(8): 1245-1262.
- Illari, P, and Williamson, J (2013) In defence of activities. *Journal for General Philosophy of Science* 44(1): 69-83.
- Joh, TH, Park, DH and, Reis, DJ (1978) Direct phosphorylation of brain tyrosine hydroxylase by cyclic AMP-dependent protein kinase: mechanism of enzyme activation. *Proceedings of the National Academy of Sciences* 75(10): 4744-4748.
- Kahn, J. (2005). Misreading race and genomics after BiDil. *Nature genetics* 37(7): 655-656.
- Kaiser, MI (2018) The components and boundaries of mechanisms. in Stuart Glennan, Phyllis Illari (eds.) *The Routledge Handbook of Mechanisms and Mechanical Philosophy*: New York pp. 116-130.
- Kaplan JM (2014) When socially determined categories make biological realities: understanding Black/White health disparities in the US. *The Monist* 93.2: 281-297
- Kaplan JM, Winther RG (2014) Realism, antirealism, and conventionalism about race. *Philosophy of Science* 81(5): 1039-1052.
- Kaufman, J. S., Dolman, L., Rushani, D., & Cooper, R. S. (2015). The contribution of genomic research to explaining racial disparities in cardiovascular disease: a systematic review. *American Journal of Epidemiology* 181(7): 464-472.

- Kauffman, SA (1971) Articulation of parts explanation in biology and the rational search for them" in Roger C. Buck and Robert S. Cohen (eds.), *PSA 1970: 257-272*: Boston Studies in the Philosophy of Science, v. 8. Dordrecht, Reidel.
- Kaup AR, Nettiksimmons J, Harris TB, Sink KM, Satterfield S, Metti AL, Ayonayon HN, Yaffe K (2015) Cognitive resilience to apolipoprotein E ϵ 4: contributing factors in black and white older adults. *JAMA Neurology* 72 (3): 340-348.
- Kennedy A, Malanowski S (2019) Mechanistic reasoning and informed consent. *Bioethics* 33(1): 162-168.
- Kitcher P (2007) Does 'race' have a future? *Philosophy and Public Affairs*: 293-317.
- Kramer MS, Ananth CV, Platt RW, Joseph KS (2006) US black vs. white disparities in fetal growth: physiological or pathological? *International Journal of Epidemiology* 35(5): 1187-1195.
- Kuzawa CW, Sweet E (2009) Epigenetics and the embodiment of race: developmental origins of US racial disparities in cardiovascular health. *American Journal of Human Biology* 21(1): 2-15.
- Kuzawa CW, Thayer Z (2011) The timescales of human adaptation: the role of epigenetic processes. *Epigenomics* 3(2): 221-234.
- Lewis, D (1986) *Philosophical Papers, Volume 2*. Philosophical Papers, Volume 2. Oxford: Oxford University Press.
- Lewontin RC (1972) The apportionment of human diversity. *Evolutionary Biology* (6): 381-398.
- Little, D (2011) Causal mechanisms and the social realm. In Illari, PM, Russo, F, and Williamson, J (eds.) *Causality in the Sciences*. Oxford: Oxford University Press.
- Logue MW, Schu M, Vardarajan BN, Farrell J, Bennett DA, Buxbaum JD, Byrd G (2014) Two rare AKAP9 variants are associated with Alzheimer's disease in African Americans. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* 10(6): 609-618.
- Love, A (2007) Functional homology and homology of function: biological concepts and philosophical consequences. *Biology & Philosophy* 22(5): 691-708.
- Lorusso L, Bacchini F (2015) A reconsideration of the role of self-identified races in epidemiology and biomedical research. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 52: 56-64.
- Machamer P (2004) Activities and causation: the metaphysics and epistemology of mechanisms. *International Studies in the Philosophy of Science* 18(1): 27-39.
- Machamer P, Darden L, Craver CF (2000) Thinking about mechanisms. *Philosophy of Science* 67(1): 1-25.

- Magnus, PD (2012) *Scientific Enquiry and Natural Kinds: From Planets to Mallards*. Houndmills: Palgrave Macmillan.
- Marks, J (2007) Long shadow of Linnaeus's human taxonomy. *Nature* 447(7140): 28-28.
- Matthews LJ (2017) On mechanistic reasoning in unexpected places: the case of population genetics. *Biology & Philosophy* 32(6): 999-1018.
- McDade TW, Ryan CP, Jones MJ, Hoke MK, Borja J, Miller GE, Kuzawa CW, Kober MS (2019) Genome-wide analysis of DNA methylation in relation to socioeconomic status during development and early adulthood. *American Journal of Physical Anthropology* 169(1): 3-11.
- McEwen BS, Gianaros PJ (2010) Central role of the brain in stress and adaptation: links to socioeconomic status, health, and disease. *Annals of the New York Academy of Sciences* 1186(1): 190-222.
- Mitchell, SD (2020) Through the Fractured Looking Glass. *Philosophy of Science* 87(5): 771-792.
- Moroz, LL (2014) The genealogy of genealogy of neurons. *Communicative & integrative biology*. 7(6): e993269.
- Müller-Wille, S. (2014). Race and history: Comments from an epistemological point of view. *Science, Technology, & Human Values* 39(4): 597-606
- Nathan, M (forthcoming) Redundant causality and robustness. CK Waters, BJ Hanley, and J Woodward (eds.) *Philosophical Perspectives on Causal Reasoning in Biology, Volume XXI of the Minnesota Studies in the Philosophy of Science*. Minneapolis: University of Minnesota Press.
- Neander, K (2002) Types of traits: function, structure and homology in the classification of traits, in André Ariew, Robert Cummins, and Mark Perlman (eds.), *Functions: New Essays in the Philosophy of Biology and Psychology*. Oxford: Oxford University Press.
- O'Donnell KJ, Meaney MJ (2017) Fetal origins of mental health: the developmental origins of health and disease hypothesis. *American Journal of Psychiatry* 174(4): 319-328.
- Öz, R, Howard, SM, Sharma, R, Törnkvist, H, Ceppi, I, Sriram, KK, ... and Westerlund, F (2020) Phosphorylated CtIP bridges DNA to promote annealing of broken ends. *Proceedings of the National Academy of Sciences* 117(35): 21403-21412.
- Parkkinen, VP, Wallmann, C, Wilde, M, Clarke, B, Illari, P, Kelly, MP, Norell, C, Russo, F, Shaw, B, Williamson, J (2018) *Evaluating Evidence of Mechanisms in Medicine: Principles and Procedures*. Springer Briefs in Philosophy. Open access: <https://link.springer.com/book/10.1007%2F978-3-319-94610-8>

- Paslaru, V (2014) The mechanistic approach of the theory of island biogeography and its current relevance. *Studies in History and Philosophy of Biological and Biomedical Sciences* 45:22-33.
- Persson, J (2010) Activity-based accounts of mechanism and the threat of polygenic effects. *Erkenntnis* 72(1): 135-149.
- Pearlman, SM, Serber, Z and, Ferrell Jr., JE (2011) A mechanism for the evolution of phosphorylation sites. *Cell* 147(4): 934-946.
- Phillips DIW, Barker DJP, Fall CHD, Seckl JR, Whorwood CB, Wood PJ, Walker BR (1998) Elevated plasma cortisol concentrations: a link between low birth weight and the insulin resistance syndrome? *The Journal of Clinical Endocrinology & Metabolism* 83(3): 757-760.
- Piel, FB, Patil, AP, Howes, RE, Nyangiri, OA, Gething, PW, Williams, TN, ... and Hay, SI (2010) Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis. *Nature communications* 1(1): 1-7.
- Pietroski, P and Rey, G (1995) When other things aren't equal: saving ceteris paribus laws from vacuity. *The British Journal for the Philosophy of Science* 46(1): 81-110.
- Potochnik, A (2010) Levels of explanation reconceived. *Philosophy of Science* 77: 59–72.
- Prince M, Bryce R, Wimo AE, Ribeiro W, Ferri CP (2013) The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* 9(1): 63-75.
- Puckrein, G. (2006). BiDil: From Another Vantage Point: Despite criticisms of BiDil, its clinical benefits for African Americans have held up to scientific scrutiny. *Health Affairs* 25(Suppl1): W368-W374.
- Rao, S, Segar, MW, Bress, AP, Arora, P, Vongpatanasin, W, Agusala, V, ... and Pandey, A (2020) Association of genetic West African ancestry, blood pressure response to therapy, and cardiovascular risk among self-reported Black individuals in the Systolic Blood Pressure Reduction Intervention Trial (SPRINT). *JAMA cardiology* PMID: 33185651.
- Reimers, MA, Craver, C, Dozmorov, M, Bacanu, SA, and Kendler, KS (2019) The coherence problem: Finding meaning in GWAS complexity. *Behavior genetics* 49(2): 187-195.
- Reitz C, Jun G, Naj A, Rajbhandary R, Vardarajan BN, Wang LS, Valladares, O et al. (2013) Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E ϵ 4, and the risk of late-onset Alzheimer disease in African Americans. *JAMA* 309(14): 1483-1492.
- Root M (2003) The use of race in medicine as a proxy for genetic differences. *Philosophy of Science* 70(5): 1173–1183.

- Rosenberg, NA, Pritchard, JK, Weber, JL, Cann, HM, Kidd, KK, Zhivotovsky, LA, and Feldman, MW (2002) Genetic structure of human populations. *Science* 298(5602): 2381-2385.
- Russo F, and Williamson J (2007) Interpreting causality in the health sciences. *International Studies in the Philosophy of Science* 21(2): 157-170.
- Safieh M, Korczyn AD, Michaelson DM (2019) ApoE4: an emerging therapeutic target for Alzheimer's disease. *BMC Medicine* 17(1): 64.
- Scholaske L, Buss C, Wadhwa PD, Entringer S (2018). Acculturation and interleukin (IL)-6 concentrations across pregnancy among Mexican-American women. *Brain, Behavior, and Immunity* 73: 731-735.
- Sedivy, R, Windischberger, C, Svozil, K, Moser, E, and Breitenecker, G (1999) Fractal analysis: an objective method for identifying atypical nuclei in dysplastic lesions of the cervix uteri. *Gynecologic Oncology* 75(1): 78-83.
- Sesardic N (2010) Race: a social destruction of a biological concept. *Biology & Philosophy* 25(2): 143-162.
- Sider, T (2013) *Writing the Book of the World*. Oxford: Oxford University Press.
- Spencer QNJ (2015) Philosophy of race meets population genetics. *Studies in History and Philosophy of Biological and Biomedical Sciences* 52: 46–55.
- Spencer QNJ (2018) A racial classification for medical genetics. *Philosophical Studies* 175(5): 1013-1037.
- Steinman, L (2001) Multiple sclerosis: a two-stage disease. *Nature Immunology* 2(9): 762-764.
- Stepler, KE, Robinson RA (2019) The potential of 'omics to link lipid metabolism and genetic and comorbidity risk factors of Alzheimer's Disease in African Americans. In *Reviews on Biomarker Studies in Psychiatric and Neurodegenerative Disorders* (pp. 1-28). Springer, Cham.
- St-Pierre J, Hivert MF, Perron P, Poirier P, Guay SP, Brisson D, Bouchard L (2012) IGF2 DNA methylation is a modulator of newborn's fetal growth and development. *Epigenetics* 7(10): 1125-1132.
- Sternthal MJ, Slope N, Williams DR (2011) Racial disparities in health. *Du Bois Review: Social Science Research on Race* 8(01): 95-113.
- Stewart CP, Oaks BM, Laugero KD, Ashorn U, Harjunmaa U, Kumwenda C, Dewey KG (2015) Maternal cortisol and stress are associated with birth outcomes, but are not

- affected by lipid-based nutrient supplements during pregnancy: an analysis of data from a randomized controlled trial in rural Malawi. *BMC Pregnancy and Childbirth* 15(1): 346.
- Strevens, M (2004) The causal and unification approaches to explanation unified—causally. *Noûs* 38(1): 154-176.
- Strevens, M (2008) *Depth: An Account of Scientific Explanation*. Cambridge, MA: Harvard University Press.
- Strevens, M (forthcoming) The causal role of aggregate properties. CK Waters, BJ Hanley, and J Woodward (eds) *Philosophical Perspectives on Causal Reasoning in Biology, Volume XXI of the Minnesota Studies in the Philosophy of Science*. Minneapolis: University of Minnesota Press.
- Stevano, S (2019) The randomistas won the 2019 Nobel Prize for Economics. Will they also win the debate on development economics? *Economics* 12: 20.
- Sullivan S (2013) Inheriting racist disparities in health: Epigenetics and the transgenerational effects of white racism. *Critical Philosophy of Race* 1(2): 190-218.
- Tabery J (2014) *Beyond versus: The struggle to understand the interaction of nature and nurture*. MIT Press, Boston.
- Tam, V, Patel, N, Turcotte, M, Bossé, Y, Paré, G, and Meyre, D (2019) Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics* 20(8): 467-484.
- Tarver-Carr ME, Powe NR, Eberhardt MS, LaVeist TA, Kington RS, Coresh J, Brancati FL (2002) Excess risk of chronic kidney disease among African-American versus white subjects in the United States: a population-based study of potential explanatory factors. *Journal of the American Society of Nephrology* 13(9): 2363-2370.
- Taylor P (2011) Rehabilitating a biological notion of race? A response to Sesardic. *Biology & Philosophy* 26(3): 469-473.
- Temple R, Stockbridge NL (2007) BiDiI for heart failure in black patients: The US Food and Drug Administration perspective. *Annals of Internal Medicine* 146(1): 57-62.
- Thagard, P (1989) Explanatory coherence. *Behavioral and Brain Sciences* 12(3): 435-502.
- Trogon, K (2018) Grounding-mechanical explanation. *Philosophical Studies* 175(6): 1289-1309.
- Valles SA (2012) Heterogeneity of risk within racial groups, a challenge for public health programs. *Preventive Medicine* 55(5): 405-408.
- Vlietinck R, Derom R, Neale MC, Maes H, Van Loon H, Derom C, Thiery M (1989) Genetic and environmental variation in the birth weight of twins. *Behavior Genetics* 19(1): 151-161.

- Waters, CK (2007) Causes that make a difference. *The Journal of Philosophy* 104(11): 551-579.
- Weslake, B (2010) Explanatory depth. *Philosophy of Science* 77(2): 273-294.
- Whitfield JB, Treloar SA, Zhu G, Martin NG (2001) Genetic and non-genetic factors affecting birth-weight and adult body mass index. *Twin Research* 4(05): 365-370.
- Wilson V, Rogers III WM (2016) Black white wage gaps expand with rising wage inequality. *Economic Policy Institute*. Web. <<http://www.epi.org/publication/black-white-wage-gaps-expand-with-rising-wage-inequality/#epi-toc-7>>
- Wimsatt, William C 1981 Robustness, reliability, and overdetermination, in M. Brewer and B. Collins (eds.), *Scientific Inquiry and the Social Sciences*. San Francisco, CA: Jossey-Bass, pp. 124-163.
- Woodward, J (1997) Explanation, invariance, and intervention. *Philosophy of Science* 64: S26-S41.
- Woodward, J (2003) *Making Things Happen: A Causal Theory of Explanation*. New York: Oxford University Press.
- Woodward, J (2010) Causation in biology: stability, specificity, and the choice of levels of explanation. *Biology and Philosophy* 25: 287-318.
- Ylikoski, P and Kuorikoski, J (2010) Dissecting explanatory power. *Philosophical Studies* 148:201-219.
- Zakharia F, Basu A, Absher D, Assimes TL, Go AS, Hlatky MA, Sidney S (2009) Characterizing the admixed African ancestry of African Americans. *Genome biology* 10(12): R141.
- Zijlmans MA, Riksen-Walraven JM, de Weerth C (2015) Associations between maternal prenatal cortisol concentrations and child outcomes: a systematic review. *Neuroscience & Biobehavioral Reviews* 53: 1-24.