

2020

# Back to the events themselves: on what events are and how we perceive them

---

<https://hdl.handle.net/2144/39869>

*Boston University*

BOSTON UNIVERSITY  
GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

**BACK TO THE EVENTS THEMSELVES:  
ON WHAT EVENTS ARE AND HOW WE PERCEIVE THEM**

by

**KUEI-CHEN CHEN**

B.B.A., National Taiwan University, 2009  
M.A., The University of Chicago, 2011

Submitted in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

2020



Approved by

First Reader

---

Walter Hopp, Ph.D.  
Associate Professor of Philosophy

Second Reader

---

Daniel O. Dahlstrom, Ph.D.  
John R. Silber Professor of Philosophy

Third Reader

---

Jeffrey Yoshimi, Ph.D.  
Associate Professor of Philosophy and Cognitive Science  
University of California, Merced

Fourth Reader

---

Marc Gasser-Wingate, Ph.D.  
Assistant Professor of Philosophy

“Meanings inspired only by remote, confused, inauthentic intuitions – if by any intuitions at all – are not enough: we must go back to the ‘things themselves’”.

Edmund Husserl, *Logical Investigations* ([1913a] 1970, 1:168)

## DEDICATION

To my family,  
for being there for me every step of the way

## ACKNOWLEDGMENTS

The first person I would like to thank is Walter Hopp, my advisor. Walter showed me how to do philosophy through both passionate teaching and exemplary work; it is hard to describe the intellectual excitement I felt during my conversations with him. His care for students, open-mindedness and endless phenomenological insights made the graduate school one of the most fulfilling experiences in my life. I am also deeply indebted to the other members of my dissertation committee. Dr. Daniel Dahlstrom helped me to see things from a broader and more historically grounded perspective. Jeff Yoshimi read my drafts at various points and gave me extensive feedback. Our exchanges have made a tremendous impact on my philosophical views. Marc Gasser-Wingate kindly chaired my defense and asked perceptive questions that led me to re-examine some of my assumptions in the dissertation. In addition to my committee members, I would like to thank Dr. Lee-Chun Lo, for introducing me to phenomenology and for the continued guidance after I graduated from college.

I am grateful to my friends for their company over these years. Colin Cmiel was my roommate for six years, who filled our apartment with fun conversations. Guy Schuh has always been a friend in need. Maité Cruz Tleugabulova taught me many wonderful things both in and outside philosophy, including the secret about tomatoes. Qiong Wu's philosophical orientation is quite different from mine, but our conversations on metaphysics were a highlight of my time in the graduate school. I-Kai Jeng and Robin Weiss hosted the best parties I had in Boston. Yen-Hsiang Wang shared countless fascinating ideas with me in our reading group. Shih-Shiang Hung and Yu-Cheng Wu were the go-to people when I

wanted to hang out with someone; many of my happiest moments over these years were shared with them.

My family is with me in times good and bad. My parents have always been supportive and understanding. It would have been impossible for me to finish the dissertation if I didn't have the encouragement from them. And I have the best kind of brother one could ask for, a fellow traveler in the academic journey and a trusted friend in life.

Finally, I would like to thank Yi-Jing – for showing me how every single day can be a blessing.



**BACK TO THE EVENTS THEMSELVES:  
ON WHAT EVENTS ARE AND HOW WE PERCEIVE THEM  
KUEI-CHEN CHEN**

Boston University Graduate School of Arts and Sciences, 2020

Major Professor: Walter Hopp, Associate Professor of Philosophy

**ABSTRACT**

A task confronting all the theoretical branches of philosophy is Peacocke's *Integration Challenge*: "providing, for a given area, a simultaneously acceptable metaphysics and epistemology, and showing them to be so" (1999, 1). When it comes to the study of everyday events such as basketball shots and ripples on a pond, there is a phenomenological analogue of the Integration Challenge: the twofold task of explicating both the nature of such events and the way they show up in our perceptual consciousness. In this dissertation, I propose an account of events in response to the twofold challenge. On the one hand, an event is a complex entity constituted from tropes in accordance with its kind. On the other hand, our perceptual experiences of events differ from our other varieties of experiences because they uniquely feature the awareness of certain temporal phenomena that function as the boundaries between events.

The dissertation is divided into five chapters. After a short introductory chapter, I develop a metaphysical theory of events. I follow a framework based on Evnine (2016) and divide theories of events into two groups: those that invoke the constitution relation and those that do not. Chapter 2 reviews and argues against the major theories in the first group, whereas Chapter 3 defends my own view against several alternatives in the second group.

Roughly, my view is that a number of events jointly constitute another if and only if the event-kind the latter falls under makes the latter dependent upon the former.

The remaining chapters defend a Husserlian view about event perception. Chapter 4 situates my favored view in the theoretical landscape. I argue for a representational view by drawing on Husserlian ideas, in particular the idea that any perceptual content has an expectational component in addition to a component that represents what is strictly visible. Chapter 5 expands on the view and supplies an analysis of event perception. By combining the theory of perceptually based expectations proposed by Yoshimi (2016) with results from linguistics and psychology, I argue that event perception can be better understood with the theoretical apparatus of possible worlds.

## TABLE OF CONTENTS

|   |      |
|---|------|
| DEDICATION.....                                       | v    |
| ACKNOWLEDGMENTS .....                                 | vi   |
| ABSTRACT .....  | viii |
| TABLE OF CONTENTS.....                                | x    |
| LIST OF FIGURES .....                                 | xiii |
| LIST OF ABBREVIATIONS .....                           | xiv  |
| CHAPTER 1 INTRODUCTION .....                          | 1    |
| 1.1    Motivating the Project.....                    | 1    |
| 1.2    An Overview of Main Arguments .....            | 6    |
| CHAPTER 2 EVENTS: NON-HYLOMORPHIC ACCOUNTS .....      | 10   |
| 2.1    Preliminaries.....                             | 12   |
| 2.1.1    “Events”: Delimiting the Subject Matter..... | 12   |
| 2.1.2    Existence and Identity.....                  | 19   |
| 2.2    The Space-Time View .....                      | 20   |
| 2.3    The State-of-Affairs View.....                 | 26   |
| 2.4    The Instantiation View.....                    | 36   |
| 2.4.1    Kim’s Account.....                           | 36   |
| 2.4.2    Cleland’s Account .....                      | 39   |
| 2.4.3    Bennett’s Account.....                       | 47   |

|  |  |     |
|--|--|-----|
| 2.5  | Conclusion .....                             | 57  |
| CHAPTER 3 EVENTS: A HYLOMORPHIC PROPOSAL .....                 |  | 58  |
| 3.1  | Constitution in the Domain of Events .....   | 59  |
| 3.2  | Crowther’s Account.....                      | 61  |
| 3.3  | Evnine’s Account.....                        | 72  |
| 3.4  | Jones’ Account.....                          | 80  |
| 3.5  | An Alternative Hylomorphic Account.....      | 88  |
| 3.5.1  | Fine and Johnston on Event Constitution..... | 88  |
| 3.5.2  | Dependence-Based Hylomorphism .....          | 91  |
| 3.6  | Objections and Replies.....                  | 102 |
| 3.6.1  | Realization, Not Constitution .....          | 102 |
| 3.6.2  | Events Have No Contingent Properties .....   | 105 |
| 3.7  | Conclusion .....                             | 113 |
| CHAPTER 4 PERCEPTION: A HUSSERLIAN REPRESENTATIONAL VIEW ..... |  | 114 |
| 4.1  | Vehicles, Contents and Objects.....          | 117 |
| 4.2  | Two Common Views .....                       | 125 |
| 4.2.1  | Propositionalism.....                        | 125 |
| 4.2.2  | Pictorialism.....                            | 131 |
| 4.3  | The Husserlian Dual-Component View .....     | 136 |
| 4.3.1  | Husserl on Perception.....                   | 136 |
| 4.3.2  | HDC: A First Formulation.....                | 139 |
| 4.3.3  | How Many Kinds of Contents Are There?.....   | 147 |

|  |  |     |
|--|--|-----|
| 4.3.4  | Are Experiences Object-Involving?                          | 151 |
| 4.3.5  | HDC: A Second Formulation                                  | 160 |
| 4.4  | Challenges from the Relational View                        | 175 |
| 4.5  | Conclusion   | 180 |
| CHAPTER 5 PERCEPTION: THE EXPERIENCE OF EVENTS |  | 181 |
| 5.1  | Why Care about Event Perception?                           | 184 |
| 5.2  | Contrast Case 1: The Experience of Individuals             | 187 |
| 5.3  | Contrast Case 2: The Experience of Time                    | 193 |
| 5.4  | The Experience of Events                                   | 206 |
| 5.4.1  | Experiencing Completing Boundaries                         | 206 |
| 5.4.2  | An Excursion into Natural Language Semantics               | 211 |
| 5.4.3  | Specifying Expectational Contents: The Expectational Base  | 217 |
| 5.4.4  | Specifying Expectational Contents: The Salience Constraint | 225 |
| 5.4.5  | BRV Restated   | 237 |
| 5.5  | Revisiting the Central Thought Experiment                  | 246 |
| 5.5.1  | Are All Expectational Contents Non-Perceptual Contents?    | 246 |
| 5.5.2  | Are All Expectational Contents Belief Contents?            | 250 |
| 5.6  | Conclusion   | 256 |
| BIBLIOGRAPHY                                   |  | 258 |
| VITA   |  | 267 |

## LIST OF FIGURES

|  |     |
|--|-----|
| Figure 1. The London Eye.....  | 133 |
| Figure 2. Point-light displays: an example .....   | 191 |
| Figure 3. The experienced instances of the specious present and the experienced units of the<br>perceived event..... | 199 |

## LIST OF ABBREVIATIONS

|           |                                     |
|-----------|-------------------------------------|
| BRV ..... | Boundary Representation View        |
| CP .....  | Constitution by a Plurality         |
| DBH ..... | Dependence-Based Hylomorphism       |
| EAT ..... | Event Perception as Time Perception |
| HDC ..... | Husserlian Dual-Component View      |
| IPL ..... | Initial Phase of a Lob              |
| PA .....  | Propositional-Attitude View         |
| UT .....  | Unified Theory                      |

## CHAPTER 1

### INTRODUCTION

This dissertation addresses two issues at the intersection of metaphysics, philosophy of mind and phenomenology: what is the nature of events, and what distinguishes our conscious perception of events from the other forms of conscious perception? I defend a hylomorphic view in response to the first issue and a Husserlian view in response to the second. On the one hand, to be an event is to consist of certain form and matter. On the other hand, what makes conscious event perception special is that in undergoing them, we anticipate temporal boundaries in a distinctively perceptual manner. As a whole, my dissertation represents my attempt to pursue philosophy in a way that respects the divergence between different traditions but makes the most of their points of convergence.

#### 1.1 Motivating the Project

Why care about events? Events are worth theorizing about if for no other reason than the roles they play in our everyday life. There are hardly moments in our life in which we are not participating in one event or another. Whether one is doing something as mundane as riding a bus to work or something as extraordinary as climbing Mount Everest, one is partaking in an event. In our everyday understanding of reality, therefore, events occupy a conspicuous place. We cannot make sense of the world around us without taking into account the various events that take place around us. Since philosophy is supposed to be rooted, at least in part, in common sense, the everyday significance of events alone makes events a worthy topic of philosophical investigation.



The nature of events used to be an intensely debated topic in metaphysics. Prominent metaphysicians of the era, such as David Lewis, W. V. O. Quine, Donald Davidson, Roderick Chisholm and Jaegwon Kim, among others, all contributed to the debate. As is well known, the 20th century even produced metaphysical systems that conceptualize events of some sort as the building blocks of reality, the most celebrated one being the system laid out by A. N. Whitehead in his *Process and Reality*. The theoretical landscape of contemporary metaphysics has nevertheless changed. Though the earlier theories of events are still mentioned, more often than not they are introduced for the purpose of clarifying one's presuppositions rather than for the purpose of actively developing the theories. People became less interested in the issue not because some kind of consensus had been reached on the nature of events. What happened was more aptly characterized by the words used by Stephen Kosslyn to describe a phase of the debate on the nature of mental imagery: the debate "ended with a fizzle, with no clear consensus on a resolution... most researchers apparently got tired of the issues and moved on to other topics" (Kosslyn 1994, 12). I claim that this is a prime time to revisit the debate on the nature of events and propose new theories to resolve the old disputes. This is the case because more conceptual tools are available right now. In particular, recent years have seen increasingly sophisticated attempts to adapt theories of material constitution to the analysis of events. These attempts have opened up new avenues for research, and they will be examined in detail in this dissertation.

Now, as one can tell from the title of this dissertation, my project is not exclusively metaphysical. It also addresses our perceptual experiences of events. The reason is that the kind of project I am interested in pursuing is one that may be described as *phenomenologically*

*serious metaphysics*. What this means is best illustrated with an example. Suppose I underwent a visual experience caused by a persimmon in front of me. Could I be said to veridically see that the object in front of me was a persimmon? Some might want to answer in the negative because they prioritize their theoretical commitment to physicalism. One could argue that physics is the true theory of the world. Strictly speaking, therefore, the only entities that exist are those posited by physics. Since physics doesn't posit such properties as *being a persimmon*, nothing really *is* a persimmon. When we use the phrase "is a persimmon", what we succeed in picking out is actually a pattern of activity that can be exhibited by fundamental particles. Given that persimmons don't exist, there are only two ways to make sense of my experience. First, my experience represented something veridically, but what was represented was not a persimmon. Second, my experience did represent a persimmon. But this means that it represented something that didn't exist, so it was nonveridical. Either way, I could not be said to veridically see a persimmon.

This is not how I would want to answer the question. My preferred response to the question would give more credit to the first-person perspective. To give one such response, we could begin with considering how one would make judgments about the accuracy of one's experience in everyday life. If I see something but find it missing when I take a second look, I would begin to doubt my initial experience. But nothing like the sort happened in the persimmon case. I stared at the persimmon for half a minute, and the persimmon remained in view over the entire course of my experience. Hence, if at this point I had been asked whether my initial experience had been accurate, I would have answered in the positive. Since my cognitive capacities were functioning normally and I was not in any peculiar environment, anyone theorizing about this situation should regard my judgment as providing

strong evidence for the accuracy of my experience. But insofar as my experience was accurate, the persimmon perceived by me actually existed in front of me. If we reason this way, therefore, what one can say about the metaphysics of one's perceptual object is constrained by what one would say in the phenomenological description of one's perceptual experience.

If I am really drawn to physicalism, would this approach prevent me from upholding my theoretical commitments? No. I could defend the view that there are different levels of reality connected by such relations as grounding, parthood and material constitution. If I want to endorse physicalism, all I have to say is that the fundamental level of reality consists of only the entities posited by physics. Hence, even though a phenomenologically serious approach to metaphysics puts much emphasis on first-person observations, it need not lack metaphysical rigor. Needless to say, there are many ways to pursue phenomenologically serious metaphysics. What all these approaches would have in common is the idea that one ought to be serious about how things appear in one's experience and what one's experience reveals to one. And it is worth stressing that taking the first-person perspective seriously doesn't entail commitments to such theses as phenomenalism about physical objects and the infallibility of introspection. It only entails that first-person phenomena shouldn't be easily discounted. In cases where there are good metaphysical, empirical, or phenomenological reasons for discounting certain first-person observations, they should by all means be discounted.

On this approach to metaphysics, metaphysics should be informed by phenomenology and philosophy of mind. We cannot have a satisfactory metaphysics of events if we do not, at the same time, have a satisfactory theory of our experiences of events.

This approach to metaphysics is certainly not universally accepted, but it is not idiosyncratic, either. Christopher Peacocke argues that there is a task confronting many branches of philosophy, which he calls the *Integration Challenge*: “providing, for a given area, a simultaneously acceptable metaphysics and epistemology, and showing them to be so” (1999, 1). I think the Integration Challenge needs to be met in all the major branches of theoretical philosophy. And when it comes to studying entities of the kind that we can interact with in our everyday life, such as pumpkins and storms, there is a phenomenological analogue of the Integration Challenge. We can ask: what is the nature of these everyday entities, and how is it possible for us to see things as they are when leading our everyday life? The goal of my dissertation is to defend an account of events that meets this phenomenological challenge.

At the beginning of this introductory chapter, I said that my dissertation was inspired by ideas from different philosophical traditions. Obviously, one of my goals is to engage with academic philosophy in the English-speaking world. But it is not hard to tell that a tradition that is not the most prominent in the English-speaking world has also exerted tremendous influence on my views: the phenomenological tradition, especially Husserlian phenomenology. Would the project pursued here be consistent with Husserl’s philosophical goals? I believe so. According to Husserl, part of what it is to be a physical object is to be capable of being experienced in certain characteristic ways: “We know that belonging to the essence of the thing-objects, so far as they are veritably able to be, are not only these truths or others, but that it belongs to them to be apparent in one manner or another, to show themselves in some proving acts or others...” (Husserl 2008, 434). However, Husserl also acknowledges that it is by no means an easy task to explicate how exactly the nature of a

physical object is correlated with the possible ways to experience the object. This task is so weighty that Husserl takes it to be what partly defines phenomenology: “To elucidate these connections between veritable being and knowing and so in general to investigate the correlations between act, meaning, object is the *task of transcendental phenomenology* (or transcendental philosophy)” (Husserl 2008, 434). Now, though Husserl’s point of departure in these passages is physical objects, it is clear that his conclusion applies to objects of all kinds. Insofar as there are mental acts that are experiences of events, it should be a task of phenomenology to clarify the relation between these acts, the contents they carry and the objects they represent. If so, I believe that my project adheres to Husserl’s vision for philosophy.

## 1.2 An Overview of Main Arguments

The remaining chapters of the dissertation are divided into two parts. The first part (Chapters 2 and 3) focuses on the metaphysics of events, whereas the second part (Chapters 4 and 5) defends a Husserlian view of event perception.

*Chapter 2* examines the traditional theories of events that are still influential today. These theories make different proposals about how events are individuated: by their spatiotemporal locations, by corresponding states of affairs or by the properties of which they are instances. What these theories have in common is that they don’t posit any relation similar to material constitution in the domain of events. I argue that the theories suffer from various drawbacks – some of the most serious drawbacks result from their common feature, i.e. the failure to take into account the phenomenon of constitution.

*Chapter 3* discusses several recent theories of events. All these theories pay attention to the phenomenon of constitution, even though they don't conceptualize constitution in exactly the same way. After examining the theories, I present what I take to be a better alternative to the theories. In particular, I argue that a theoretically fruitful way to make sense of events distinguishes between simple and complex events. Simple events are tropes, while complex events are bundles of simple events unified by the constitution relation. This raises a crucial question: is there a sufficiently broad conception of the constitution relation, on which events can be understood as the relata of the relation? I argue that there is. An event always falls under a certain kind, and its kind makes it dependent upon events of some other kinds. In other words, an event has needs for other events. On my proposal, if the needs of an event are satisfied by the events co-located with it, the former can be said to be constituted by the latter. A seminar can thus be said to be constituted by all the events of talking occurring in the seminar; similarly, a concert is constituted by all the performances that appear on its program. The details of my view are given in section 3.5.2; if the reader prefers to get a sense of my view before considering my arguments against alternative views, he or she could read the section first. The section presupposes materials presented in sections 2.4.3 and 3.2, but the latter sections should be independently readable.

*Chapter 4* turns to the topic of event perception. My goal is to develop a Husserlian view that satisfactorily accounts for the relation between the objects and the representational contents of perceptual experiences. I begin by reviewing some Husserlian theories of perception; after that, I propose ways to refine them. On the view I defend, perceptual experiences represent systems of tropes via two kinds of contents. Contents of the first kind represent what is

strictly visible, such as the facing side of a pumpkin, whereas contents of the second kind represent what is expected on the basis of what is strictly visible, such as the occluded side of a pumpkin. The two kinds of contents enable us to perceive particulars, but the object of an experience is not uniquely determined by its contents. As I shall argue, one of the main virtues of this view is that it explains the phenomenological difference between seeing a cluster of properties and seeing the bearer of these properties. The details of this view are laid out in sections 4.3.2 and 4.3.5.

*Chapter 5* presents my account of our perceptual experiences of events. I argue that we perceive the ending of an event, thus perceptually individuating it, when we expect an incompatible event to occur at the next moment. Such expectations count as contents of perception because they are highly perspectival in nature. If I expect to see an event of color change when I see my friend put a slice of meat into boiling water, my expectation is to see the event from my present perspective. If I were to undergo a visual experience in which I see the meat changing color from my friend's perspective, then my expectation would be violated despite the fact that an event of color change is perceived. I call these perspectival expectations "expectational contents" and propose to model the expectational content of an event-representing experience with two sets of possible worlds. The first set of worlds captures the situations that are expectable in a very broad sense: roughly put, it is the set of worlds that wouldn't surprise one given what one is now seeing and what background knowledge one has. The second set of worlds corresponds to the set of situations that are expected in a more specific sense: a world is a member of the set only if it is anticipated by one in light of the event-kind under which the event unfolding in front of one falls. With the

postulation of these two sets of worlds, the everyday notion of expectation becomes a wieldy theoretical construct. Now we are able to both disambiguate the notion and use it to spell out why the way an event is experienced depends on the extent to which one is familiar with the event. As I will try to show, the view presented in the chapter sheds light on various phenomena related to event perception, including the indeterminacy inherent in such experiences. The considerations leading to the view are somewhat complicated, however, so the reader is advised to go directly to section 5.4.1 if he or she prefers to read a relatively precise formulation of the view first.



## CHAPTER 2

### EVENTS: NON-HYLOMORPHIC ACCOUNTS

Take a 21-gun salute that consists of a series of gunshots. Can we say that the series of gunshots is an event that differs from the salute, even though their spatiotemporal locations coincide with each other? *Hylomorphism about events* is true only if we can answer the question in the positive. Though hylomorphism about events is my preferred position, non-hylomorphic theories are much more common. Before I defend my preferred position in the next chapter, therefore, I shall examine a variety of non-hylomorphic theories and argue against them in this chapter.

The definition of hylomorphism employed here comes from Simon Evnine. Evnine defines the *matter relation* to be “the relation of *being the matter of*”, and a theory is hylomorphic if and only if it meets two conditions: the theory (a) affirms the existence of entities that stand in the matter relation and (b) takes the matter relation to be irreflexive and asymmetric (2016, 3). In the example just mentioned, I take the series of gunshots to be the matter of the salute. In addition, the relation the series of gunshots bears to the salute is irreflexive because the former is not its own matter; it is also asymmetric because the salute is not the matter of the series of gunshots. According to Evnine’s definition, these make my view a version of hylomorphism – my view entails that the salute is *constituted* by the series of gunshots, just as a car is constituted by steel and rubber.<sup>1</sup>

---

<sup>1</sup> Evnine also uses the more common term “constitution relation” to refer to the matter relation, and I shall use the two terms interchangeably.

As one can tell from works reviewing the major theories of events, such as Casati and Varzi (2015) and Simons (2005), theories that do without the matter relation are probably still the mainstream in the metaphysics of events. It is therefore helpful to begin with them before one argues for hylomorphism about events. I take there to be four main non-hylomorphic approaches to events. First, *primitivism* holds that we should take events to be basic items in our ontology. The less fundamental entities need to be defined by invoking events, but we cannot define events in terms of other entities. Second, according to the *space-time view*, events are occupants of spatiotemporal regions. Take a spatiotemporal region that has the right features; whatever it contains counts as an event. We need the proviso that the regions qualified to be events have the right features, because otherwise counterexamples abound – a discontinuous region that contains an avalanche and a tornado is presumably not a single event. Third, the *state-of-affairs view* says that events are entities structurally similar to propositions. How should we understand, say, the event of Carl’s speech? The proponents of the state-of-affairs view hold that an answer can be found by considering what expressions such as “that Carl is speaking” refer to. Fourth, according to the *instantiation view*, events are entities that come into existence when properties are instantiated. For example, Carl’s speech is an event occurring now because Carl instantiates the property of speaking now.

I think the instantiation view offers the most plausible conception of events among the four approaches. Even so, it still lacks the resources to properly describe certain complex events. As I will try to show, the failure of the instantiation view is indicative of the inadequacy of non-hylomorphic theories in general. The inadequacy can only be remedied if we embrace a robust conception of constitution. Before proceeding, however, I shall note

that I will not consider primitivism in any detail. If we accept primitivism about events, not much work is left for us to do in the metaphysics of events. Are there people who actually hold primitivism? When Davidson says that his work on event identity “may be viewed as an indirect defence of events as constituting a fundamental ontological category” ([1969] 2001, 180), his remark may be interpreted as endorsing primitivism.<sup>2</sup> I don’t have much to say about primitivism. The main reason I am dissatisfied with it is that I take any form of primitivism to be the last resort. Even if events cannot be reduced to anything else, one should at least make an effort to explicate the relations between events and the other kinds of entities. It is unclear to me that the endorsement of primitivism would encourage such efforts. Therefore, I will set primitivism aside for the rest of the chapter.

This chapter is structured as follows. Section 2.1 clarifies what count as events in my discussion and briefly describes how I understand the notions of existence and identity. Sections 2.2–2.4 evaluate the space-time view, the state-of-affairs view and the instantiation view, respectively, and Section 2.5 concludes the chapter. As such, this chapter is polemical in character; my positive view will be presented in the next chapter.

## **2.1 Preliminaries**

### *2.1.1 “Events”: Delimiting the Subject Matter*

What are the events I shall be discussing in this chapter? Here are some examples: storms, laser shows, congressional hearings, my friend’s playing of tennis and the tripling of

---

<sup>2</sup> This is how Friederike Moltmann interprets Davidson (Moltmann 2003). I am not sure that Davidson has said enough to make this interpretation defensible.

a company's profit.<sup>3</sup> From this list, it should be obvious that my concern is with everyday events. Maybe the expansion of the universe or cases of quantum entanglement also count as events, but I have nothing to say about them in this chapter. My list of examples gives a rough idea of the kind of events I intend to study. However, a list is rarely sufficient for the purpose of reference-fixing. Consider a so-called "Cambridge change": when South Sudan declared independence in 2011, Kosovo became the second youngest country in the world.<sup>4</sup> But all the intrinsic properties of Kosovo remained the same; the only thing that changed was the relation Kosovo bore to the remaining countries in the world. Is Kosovo's emergence as the second youngest country an event? Should it be treated as an explanandum of our theory? A list of examples cannot tell us how we should answer the question. To remedy this problem, I shall employ the following criterion: an event is any entity that makes a Davidsonian existence claim about events true. To make sense of this criterion, let's turn to Donald Davidson's work on events.

---

<sup>3</sup> When I name an event, I will use terms like "Edmund's feeding of the cat" instead of "Edmund's feeding the cat". As Jonathan Bennett has convincingly argued, these two types of expressions exhibit important semantic differences (1988, 5–6). Compare the following sentences:

- (1) The lawyer's questioning the witness was inappropriate.
- (2) The lawyer's questioning of the witness was inappropriate.

Sentence (1) can only be read as saying "it was inappropriate for the lawyer to question the witness", but this is not the only reading sentence (2) has. Sentence (2) can also be read as saying something like "it was fine for the lawyer to question the witness, but the way the lawyer questioned the witness was inappropriate". Bennett uses the term "imperfect nominals" to denote the gerunds that appear in sentences like (1); the gerunds that appear in sentences like (2) are called "perfect nominals" (1988, 5). The referents of imperfect nominals are more fine-grained than those of perfect nominals. Given the semantic differences, it is advisable not to use the two types of nominals interchangeably.

<sup>4</sup> The idea of a Cambridge change is commonly attributed to Peter Geach (1969).

Davidson created event semantics with his 1967 paper. The main concern of the paper is the logical form we should attribute to the sentences in which action verbs appear.

Consider the following inference:

- (Premise) Carl dozed off at noon.  
 (Conclusion) Therefore, Carl dozed off.

This is obviously a valid inference, but we cannot capture its validity by adopting the textbook procedure of translating the sentences with monadic predicates. That would result in the following logical form:

- (Premise) DOZE-OFF-AT-NOON (Carl)  
 (Conclusion) DOZE-OFF (Carl)

Since DOZE-OFF-AT-NOON and DOZE-OFF are different predicates, the formalized argument is invalid.<sup>5</sup> Davidson argues that this problem can be solved by quantifying over events ([1967] 2001, 118). On his proposal, the original argument can be formalized as follows:

- (Premise)  $\exists e (\text{DOZE-OFF}(\text{Carl}, e) \wedge \text{AT}(e, \text{noon}))$   
 “There is some  $e$  such that  $e$  is a dozing off of Carl and  $e$  occurs at noon”.

---

<sup>5</sup> In this chapter, I use expressions in small capitals in two ways. First, such expressions may be used as predicates in an object language. For example, “everyone likes pizza” might be translated as “ $\forall x \text{ LIKE-PIZZA}(x)$ ”. Second, such expressions may be used in the metalanguage to denote properties and kinds. For example, I might call the property of being a rational animal “RATIONAL ANIMAL” and the kind under which all frescoes fall “FRESCO”. I shall remain neutral on the exact relation between properties and kinds.

(Conclusion)  $\exists e \text{ DOZE-OFF}(\text{Carl}, e)$

“There is some  $e$  such that  $e$  is a dozing off of Carl”.

The validity of this formalized argument follows from the rules of classical logic.

It then emerges that on the Davidsonian approach, “Carl dozed off” is in fact an existence claim about an event. The semantic value of the verb predicate “dozed off” contains an event variable,<sup>6</sup> which is bound by an existential quantifier introduced by other morphemes in the sentence.<sup>7</sup> The sentence as a whole is true if some existing entity satisfies the condition of being a dozing off of Carl. It follows that the sentence “Carl dozed off” gives us reasons to introduce an entity into our ontology: if the sentence is true, then any entity that is quantified over by the existential quantifier in the sentence should be taken to exist. Our criterion simply says that we should take all such sentences into consideration when deciding what count as existing events. The application of this criterion isn’t always easy, but the criterion provides useful constraints on ontological disputes. Consider again Kosovo’s emergence as the second youngest country; does it count as an event on our criterion? Some might say “yes”, because “Kosovo became a younger country than South Sudan” is an existence claim on the Davidsonian approach, and it is made true by Kosovo’s emergence as the second youngest country. But there is at least one other entity that appears

---

<sup>6</sup> “Semantic value” is a term of art widely used in contemporary linguistics. Alexander Miller uses the term to translate the Fregean term “*Bedeutung*” and offers the following definition: “The semantic value of any expression is that feature of it which determines whether sentences in which it occurs are true or false” (2007, 11; emphasis omitted). Miller’s reason for translating the term this way is that Fregean *Bedeutungen* are assigned to many expressions that would not typically be classified as referring expressions (2007, 11 fn. 7). For example, the *Bedeutung* of a sentence is a truth value. But if the term “refer” has the meaning it does in “the name ‘Frege’ refers to Frege”, then sentences hardly “refer” to anything. I shall follow Miller’s use of the term in this chapter.

<sup>7</sup> For details, see Kratzer (1998) and Champollion (2011).

to make the sentence true: South Sudan's declaration of independence. To decide whether the sentence has multiple truth-makers and, if not, which of the two entities makes the sentence true, we must introduce additional linguistic data and engage with theories of truth. The criterion thus points to a way to proceed when we are unsure whether an entity should be put on our list of events.

The criterion may nonetheless strike one as bizarre: what does the Davidsonian approach have to do with metaphysics? Even if the Davidsonian approach tells us something about the ontological commitments of natural language, why should we think that the approach tells us anything about what there are in the world? One reason can be found in J. L. Austin's view on ordinary language: "our common stock of words embodies all the distinctions men have found worth drawing, and the connexions they have found worth marking, in the lifetimes of many generations" ([1957] 1979, 182). Words are devices language users employ to describe their environment. If so, a common pattern of word use can only emerge if it enables language users to interact with their environment more successfully. Patterns of word use thus provide reliable, if not infallible, information about the way things are. If, as suggested by the Davidsonian approach, part of the meaning conveyed by the utterance of a verb is that certain events occur, then we should take seriously the idea that the events so picked out really are elements of reality.

Many scientifically minded philosophers would not sympathize with Austin's view. Admittedly, natural language may give us some guidance in our metaphysical pursuits. Such guidance is nevertheless vastly inferior to that offered by the natural sciences. Why should we care about what natural language tells us about, say, minds, organisms and physical objects, when we can simply heed the results of neuroscience, biology and physics? It may or

may not be true that the natural sciences have the final say about the natures of the most fundamental entities, but the sciences cannot tell us everything there is to know about the world. This point is forcefully made by J. van Brakel. In defending the view that there is no way to reduce the everyday discourse to a scientific discourse about the putative bottom-level entities, van Brakel cites schizophrenia as an example. Given that the very study of schizophrenia aims to better understand its behavioral symptoms, all theories of schizophrenia must be evaluated in terms of their behavioral implications. If so, a theory of schizophrenia can never become a purely biological theory: the biological mechanisms underlying the symptoms are only of interest to the extent that they shed light on behavioral phenomena (Van Brakel 1996, 272–73). Extrapolating from van Brakel’s claims, we could argue that the study of ordinary language tells us much about the natures of everyday entities. Indeed, it is hard to see how the natural sciences can tell us whether such entities as sovereign states, financial institutions, epistemically possible worlds and Cambridge changes should be accepted into our ontology. One could take the natural sciences to study nothing but the theoretical entities posited in them, but it then becomes unclear why the sciences should be the only authority in our attempt to understand the reality.

One could argue that the help from natural language is vacuous, because the existence and nature of everyday entities do not raise any deep metaphysical questions. One could defend such a claim on the basis of Jonathan Schaffer’s view: metaphysics, in its Aristotelian form, “does not bother asking whether properties, meanings, and numbers exist. Of course they do! The question is whether or not they are *fundamental*” (2009, 347). To show that questions about existence are not substantive, Schaffer appeals to arguments that move from claims made in everyday life to existence claims. In one of his examples, he



argues that from the true claim “Arthur Conan Doyle created Sherlock Holmes”, one can conclude that “Sherlock Holmes exists” (2009, 359). If we apply such a proof procedure to issues about events, then all kinds of events obviously exist; there is no need to study the conditions on which events occur. Does Kosovo’s emergence as the second youngest country exist? Of course. Witness the following piece of reasoning:

- (Premise 1) Kosovo became the second youngest country in 2011.
- (Premise 2) If Kosovo became the second youngest country in 2011, then Kosovo’s becoming of the second youngest country occurred in 2011.
- (Premise 3) If Kosovo’s becoming of the second youngest country occurred in 2011, then Kosovo’s becoming of the second youngest country in 2011 exists.
- (Conclusion) Therefore, Kosovo’s becoming of the second youngest country in 2011 exists.

If such arguments are sound, we should simply say that all events exist. It follows that there are no serious metaphysical questions about everyday events. Since the Davidsonian approach is concerned with everyday events, my proposal to delimit the subject matter of my metaphysical investigation with the help of the Davidsonian approach is essentially misguided.

The obvious response to such an objection is that arguments of the kind just mentioned are not valid. As Tobias Rosefeldt (2018) points out, the semantics of natural language sentences is far from transparent. Much linguistic evidence indicates a great mismatch between the class of natural language sentences that philosophers often take to be existence claims and the class of sentences that are actually existence claims. If so, even

though linguistics doesn't study the most fundamental entities, it can be of immense help in our attempt to determine the kinds of non-fundamental entities there are (Rosefeldt 2018, 176). If the above argument about Kosovo strikes one as weird, it is because the meanings of natural language terms are stretched too far in the argument. The remedy is to pay more attention to natural language semantics; only by doing so can we identify the real ontological commitments of natural language. This is why the Davidsonian approach, as a tool to clarify verb meanings, can be a valuable tool for metaphysics. It is reasonable to believe that the Davidsonian existence claims are ontologically significant; they are made true by real entities in the world. If so, a metaphysical study that takes these entities to be the subject matter is a worthwhile pursuit.

### 2.1.2 *Existence and Identity*

We have picked a criterion to single out the events that will be the subject matter of this chapter. But what are the questions about them that need to be answered? What is a metaphysical theory of events supposed to do? In his attempt to formulate such a theory, Jaegwon Kim chooses to focus on two conditions that events must satisfy: the *existence condition*, i.e. what must be the case for there to be an event, and the *identity condition*, i.e. what must be the case for two entities to count as the same event (1976, 160–61). Following Kim, one might say that the goal of a metaphysical theory of events just is to determine the existence and identity conditions for events in general. I think this is a good way to understand our tasks at hand, but the issue of identity conditions must be framed with care. As David Lewis has remarked, “There is never any problem about what makes something identical to itself; nothing can ever fail to be. And there is never any problem about what

makes two things identical; two things never can be identical” (1986d, 192–93). If so, it would seem that questions about event identity have more to do with our thoughts about events than with events themselves – aren’t we actually asking what it takes for two thoughts about events to have the same object? But if that’s the case, are we still doing metaphysics?

Jonathan Bennett points to a way out of this impasse. He argues that the real issue is to provide a *nonduplication principle*: “If  $e_1$  exactly resembles  $e_2$  in respect  $K$ , then  $e_1$  is  $e_2$ ” (1988, 96).<sup>8</sup> Consider an analogy with meanings. How are meanings individuated? A possible answer is that they are individuated by the sentences that can express them. This suggests a nonduplication principle for meanings: if  $m_1$  and  $m_2$  exactly resemble each other with respect to the sentences that can express them, then  $m_1$  and  $m_2$  are the same meaning. Such a principle enables us to draw the inference from partial qualitative identity to numerical identity. It is weaker than the so-called principle of the identity of indiscernibles, because it doesn’t require us to consider every single property. Bennett’s idea is that the search for the conditions of event identity should be construed as the search for the nonduplication principle applicable to events. This is how I will understand the issue of identity in what follows.

## 2.2 The Space-Time View

Let’s now examine the space-time view. As we have seen, the view takes events to be what are contained in spatiotemporal regions. Two prominent defenders of the space-time

---

<sup>8</sup> Bennett says that “nonduplication principle” is David Lewis’ term, but he doesn’t say where Lewis actually uses the term.

view are W. V. O. Quine and David Lewis. What are the existence and identity conditions posited by their views? Quine takes events to be broadly construed physical objects, where such an object “is the material content of any portion of space-time, however small, large, irregular, or discontinuous” (1985, 167). It follows that there is an event if and only if there is a non-empty spatiotemporal region. Quine also offers a straightforward criterion of event identity: two events are numerically distinct if and only if they occupy different regions (1985, 167). Therefore, if a showing of the movie *Arrival* has the same spatiotemporal location as a physical event in which light is projected on a screen, the showing is identical to the physical event. I find this to be a fairly unattractive view. On the assumption that the showing and the physical event of light projection both exist, they can be said to have different properties. The showing is an expression of Denis Villeneuve’s artistic intentions; the physical event is not. It follows that the showing couldn’t take place if Villeneuve didn’t exist, but the physical event certainly could. This means that the two events have different aesthetic and modal properties, but Quine’s view cannot accommodate their difference.

Lewis’ view can be regarded as an improvement over Quine’s. Lewis takes events to be classes that contain actual and possible spatiotemporal regions as their members (1986c, 245). Suppose the aforementioned showing of *Arrival* attracted an audience of 100 people. The numerically identical showing would still have occurred, however, had there only been 99 people in the audience. Let  $r$  and  $r^*$  denote the spatiotemporal regions in which the actual and possible showings occur. While  $r$  is a part of the actual world,  $r^*$  is a part of a merely possible world. On the simplifying assumption that the showing only exists in these two worlds, Lewis’ view implies that the showing just is the class whose only members are  $r$  and  $r^*$ . Such a view avoids the problem facing Quine’s view. While the physical event of light

projection exists in those worlds where Villeneuve has never been born, the showing does not. On Lewis' view, it follows that the two events are to be identified with different classes of regions; his view thus supports the observation that the two events are numerically distinct. Despite its virtues, nevertheless, Lewis' view also has its quirks. To begin, it may strike one as bizarre to identify events with spatiotemporal regions or classes thereof. Since a region need not contain what it actually contains, regions are presumably distinct from their occupants. If so, wouldn't it be a category mistake to identify events with regions, as opposed to their occupants? This is a legitimate worry, but it shouldn't be a serious problem on a charitable interpretation of Lewis. Lewis seems to use the term "region" loosely so that its referents could include both spatiotemporal regions and their occupants. It is also an option available to him to make spatiotemporal regions the most fundamental property-bearers; I don't know whether Lewis would be willing to do so, but his talk of the "qualitative character" of a region at least suggests this possibility (1986c, 262). If he takes this route, then the occupants of a region can certainly be thought of as the qualities of the region. This is not the only worry one might have about Lewis' view, however. As I see it, Lewis' view suffers from a more serious problem, identified by Evinine in a different context. To explain this problem, a brief digression is in order.

In reviewing theories of material constitution, Evinine reconstructs the Modal Cross-Temporal Calculus of Individuals (MCCI) developed by Judith Jarvis Thomson (1983, 1998) and interprets her view as committed to a somewhat set-theoretic understanding of physical objects (Evinine 2016, 37–40). Take a physical object and the set of entities that bear the parthood relation to the object; Thomson's view implies that the identity of the object is determined by the members of this set. To state the view more precisely, let's consider the

notion of fusion in the version of MCCI presented by Evnine. Let  $\mathcal{A}$  be a class of physical entities. Then one of the axioms of MCCI guarantees that for any  $a \in \mathcal{A}$  such that  $a$  exists at time  $t$  in possible world  $w$ , there is an entity  $x$  such that  $x$  shares at least one part with  $a$  at  $t$  in  $w$ . In other words, given any class of physical objects, there is an entity that is the fusion of these objects exactly when these objects exist and in exactly those worlds where the objects exist. According to Evnine, this is the theoretical apparatus used by Thomson to model the relation a physical object bears to its ever-changing parts. Take, for example, my copy of *Monadology*; call it  $M$ . Let  $P$  be the collection of  $M$ 's pages and all the other physical components of  $M$ , such as its front cover. In the actual world,  $M$  just is the fusion of  $P$ 's members. This is nevertheless something contingent. If  $M$  were to miss a blank page, it would still exist;  $M$  thus exists in many possible worlds where  $M$ 's parts are not the members of  $P$  but those of another collection. As a result,  $M$  cannot be straightforwardly identified with any single fusion of parts;  $M$  must be understood as a class of world-fusion pairs or fusions relative to worlds. According to Evnine's reconstruction of Thomson's view, a physical object is nothing over and above such a class.

Evnine argues that this conception of physical objects is problematic. His main objection is that such a conception lacks the resources to differentiate between ordinary objects and weird fusions. Consider London Brain, an entity whose only parts are the London Eye and my brain. It can certainly be named, but there is nothing that unifies its parts other than my inexplicable impulse to use it as an example. According to Evnine, it would be a mistake to model ordinary objects upon fusions like London Brain. Doing so would make the individuality of objects dependent upon human activity: "It will be merely a contingent feature of our interests (or possibly a deeper but still contingent feature of the

way our minds work) which such fusions we pick out and name” (Evnine 2016, 40). Given that this is implied by Thomson’s view, Thomson’s view should be rejected.

I think Lewis’ view of events is susceptible to a similar criticism. Applying Evnine’s insight to the metaphysics of events, we should make it a desideratum for any satisfactory theory of events to answer the following question: given a real event, what makes it a single event, as opposed to an arbitrary collection of entities? That is, what is the principle of unity that governs events? I don’t think Lewis has the resources to answer this question. Just as Thomson’s view lumps together real objects and London Brain, Lewis’ view treats real events and arbitrary collections of regions on a par. To see this, consider  $MF$ , the class of all the regions in which I type on a keyboard at home now while a monkey eats bananas on Mount Fuji. Let’s assume that all the members of  $MF$  exist at the same time.  $MF$  is presumably not a single event; nothing unifies the various entities that are involved. By contrast, the event that is my typing now is a real event. What can Lewis say to differentiate  $MF$  from my typing now? Lewis does hold that not every class of regions counts as an event. Certain conditions must be met for a class of regions to be *formally eligible* to be an event: such a class must contain “one region each from some worlds, none from others, and never more than one from the same world” (1986c, 245). However, the difference between  $MF$  and my typing cannot be one regarding formal eligibility;  $MF$  meets the criterion of formal eligibility. There may be many monkeys eating bananas on Mount Fuji, but we could simply select one such monkey from each possible world in deciding the members of  $MF$ .

What else can Lewis say? He does say that the members of an event should not be too dissimilar: the classes qualified to be events are those “consisting of regions of many worlds united by suitable relations of similarity” (1986c, 247). We don’t have to delve into

the details of Lewis' view about similarity to see that the members of  $MF$  are indeed very dissimilar. The description of  $MF$  doesn't pick out any single monkey, nor does it specify the number of bananas that the monkey has to eat. Hence, a region where a tiny monkey eats one banana and a region where King Kong eats a ton both belong to  $MF$ ; these spatiotemporal regions don't share many features with each other. Lewis could then invoke his requirement on similarity to rule out the eventhood of  $MF$ . If so, one could argue that  $MF$  is not a counterexample to Lewis' account. However, this response doesn't really solve the problem. Lewis still hasn't told us what it is that unifies a real event but is missing in  $MF$ . Similarity is insufficient for unity. Consider an analogous case in the domain of people. Let  $W$  be the class of all the possible worlds in which I have a doppelgänger. Take the class that contains the actual me and my doppelgängers in the members of  $W$ . Can this class be identified with a single person? Since neither I nor my doppelgänger can be identified with this class, the answer is "no". This is so despite the fact that the members of this class are perfectly similar. It follows that unity cannot be confounded with similarity. However similar the members of a class are, the class might still lack unity. If so, Lewis' attempt to reduce events to classes of regions is fundamentally flawed.

As a rejoinder, a proponent of Lewis' view could argue that there is no need to derive any principle of unity from Lewis' similarity requirement. If certain spatiotemporal regions belong to the same class, they already feature a kind of unity: the joint membership in a class. Lewis' events obviously feature such minimal unity. However, this is a very weak kind of unity; many would take any two things to form a class. Classes do not even obey the axioms of set theory, as witnessed by the standard mathematical definition of a proper class: "A *proper class* is a class which is not a set" (Di Nasso 2002, 316). Even if joint membership



in a class counts as some kind of unity, it falls short of the kind of unity featured by events. The appeal to classes thus fails to show that Lewis' events are sufficiently different from London Brain. We can certainly name the Lewisian events and postulate certain conditions for their existence, but these names hardly pick out any genuine elements of reality.

To be sure, Lewis would not be discouraged by my objections, as he has a vision for metaphysics that is very different from mine. His famous doctrine of Humean supervenience states that "if two worlds match perfectly in all matters of particular fact, they match perfectly in all other ways too" (1986a, 111). If the goal of metaphysics is to defend such supervenience, then we would have said enough about any event once we have specified all the spatiotemporal regions at which the event could possibly occur. The event certainly supervenes on this class of regions, whether or not this class of regions features any form of unity. It could be that all my dissatisfaction with Lewis' theory boils down to my view that such a vision for metaphysics is objectionable. Of course, whether or not the idea of Humean supervenience is tenable is not an issue I could possibly address here. My goal right now is only to show that the space-time view of events carries a heavy theoretical burden. If one is unwilling to undertake a metaphysical project that is Lewisian in spirit, one would do better to look elsewhere.

### **2.3 The State-of-Affairs View**

Let's now turn to the state-of-affairs view. Though it is unclear whether ordinary English hypostatizes the referent of the term "state of affairs", the term is widely used in everyday contexts. For example, a Democrat might say "the U.S. government was shut down and thousands of federal workers didn't get paid; that was an extremely unfortunate state of

affairs”. The phrase “an extremely unfortunate state of affairs” in this sentence is a nominal predicate that describes a government shutdown, which appears to be an event. Uses of the term like this might provide *prima facie* reasons to identify events with states of affairs. In the philosophical literature on events, the champion of the state-of-affairs view is Roderick Chisholm. Therefore, I shall begin by examining his version of the view.

Chisholm defines states of affairs as broadly construed propositions, which he in turn identifies with the possible objects of propositional attitudes (1970, 19–20). In doing so, he seems to presuppose a widely accepted view about belief-like states, which may be called the *propositional-attitude view* (PA). According to PA, mental states like belief and doubt are attitudes one could take towards propositions. Suppose I believe that Edmund is playing tennis. PA would characterize my belief as an affirmative attitude I take towards the proposition ⟨Edmund is playing tennis⟩. This proposition is then the object of my belief, which means that it is a state of affairs on Chisholm’s view. According to Chisholm, a state of affairs is either an event or a narrowly construed proposition: while the latter is “any state of affairs which is necessarily such that either it or its negation does not occur”, the former is “any contingent state of affairs which is not a proposition and which implies change” (1970, 20). It follows that the sentence “Edmund is playing tennis” expresses an event: it is possible for Edmund’s playing tennis to occur, so is Edmund’s not playing tennis.

Chisholm’s main argument for his view seems to be the following: events are capable of recurring, and the state-of-affairs view best accommodates this observation (1970, 15). Suppose my car accelerates now, moves at a constant velocity for 10 seconds and then accelerates again. According to Chisholm’s view, a single event that is the acceleration of my car occurs over two disjoint time intervals separated by 10 seconds. The acceleration event is

thus capable of existing during the 10 seconds when it doesn't occur; otherwise, upon recurring, the event would have to regain its identity after ceasing to be (1970, 23). How is this possible? The state-of-affairs view offers a simple explanation: the state of affairs that my car accelerates obtains for a while, stops obtaining for 10 seconds and then obtains again. Just as a false proposition still exists as a proposition, a non-obtaining state of affairs may still exist. If so, recurrence can be understood as a change in a state of affairs that outlasts the change.

What should we say about Chisholm's theory? As I see it, Chisholm is to be credited with the insight that there may be important analogies between the occurrence of an event and the obtaining of a state of affairs. There may be ways to defend the state-of-affairs view by further pursuing these analogies, as I shall soon discuss. That said, I think Chisholm's own version of the state-of-affairs view is bizarre. Why should we think that a numerically identical event could occur on two separate occasions? In our car example, isn't it just simpler to say that a second event occurs when my car accelerates again? There might be larger theoretical concerns that motivate Chisholm's view, but I cannot find any explicit statements about such concerns. One way to interpret Chisholm charitably is to hypothesize that Chisholm is actually theorizing about event-types rather than event-tokens. Unfortunately, this interpretive hypothesis is easily refuted: Chisholm's proposal is obviously intended to be an alternative to a theory that posits both types and tokens (1970, 15). The best one could do to defend Chisholm, therefore, is to claim that it is fine for a philosophical theory to be counterintuitive, insofar as it is not objectionable on independent grounds.

But Chisholm's theory *is* objectionable on independent grounds. To begin, many events seem incapable of recurring – the big bang and my birth, for example. These events

not only fail to recur in the actual world but also fail to do so in every other possible world. It doesn't make much sense to say "the big bang still exists now; it just no longer occurs". Why, then, should we take these events to be non-obtaining states of affairs that nonetheless continue to exist? An even more serious problem with Chisholm's view is that it presupposes PA. It is certainly not uncommon for philosophers to consider belief-like states as directed at propositions, but I think PA is misguided. Suppose I believe that the brick in front of me is red. What is the object of my belief? The simplest answer is: the red brick in front of me. This answer seems good enough, and it doesn't commit us to the existence of anything proposition-like. The friend of PA might argue that the simple answer doesn't account for incorrect beliefs. Suppose I am mistaken about the red brick and believe that it is blue. Since nothing red figures in my belief, its object cannot be the red brick. Rather, the object of my belief is the proposition or state of affairs that the brick in front of me is blue. This line of reasoning is problematic, because it makes the incorrectness of my belief inexplicable. Why is it incorrect to believe that the brick is blue, if what is expressed by "that the brick is blue" is precisely the object of my belief? Doesn't my belief represent this object in a perfectly accurate way? The better way to analyze the case, I submit, is to say that the *content* of my belief attributes the wrong properties to the *object* of my belief. The content of my belief can be expressed by the sentence "the brick in front of me is red"; my belief is incorrect because its object, the brick, is not red.<sup>9</sup> There is no need to invoke proposition-like entities to explain the incorrectness of my belief.

---

<sup>9</sup> For more discussion of the content-object distinction, see section 4.1.

As a rejoinder, the friend of PA could argue that the incorrectness of my belief can be easily explained by their view as well: my belief is incorrect because the object of my belief doesn't correspond to anything in reality. But isn't a proposition, as conceptualized by PA, a part of reality? Maybe the friend of PA would say that the proposition or state of affairs represented by my belief doesn't obtain. My belief is incorrect for exactly the reason that it purports to represent something that fails to obtain. I am not satisfied with this response, but instead of directly arguing against it, I would only point out that this response has undesirable implications. If the object of an incorrect belief is a non-obtaining state of affairs, the object of a correct belief is presumably an obtaining state of affairs. It follows that the red brick is not the object of my belief; it is at best an object of my belief in a derivative sense. But why cannot my belief, as it were, directly reach the brick? Why does my representation of the brick have to be mediated by the representation of a proposition-like entity? The friend of PA could argue that I am prejudiced against states of affairs. The brick is a part of a state of affairs, and the parthood relation the brick bears to the state of affairs is the same relation that holds between, say, a brick and a wall. If it is misleading to say that the representation of a wall is mediated by the representations of its constituent bricks, it is no less misleading to say that the representation of a state of affairs has to be mediated by the representations of its parts. A problem with this response is that the parthood relation in these two cases cannot be the same. Bricks and walls belong to the same ontological category: they are both physical objects. A state of affairs is nonetheless not a physical object. While the parthood relation between a brick and a wall can be defined within a single category, i.e. the category of physical objects, the parthood relation between a brick and a state of affairs has to bridge two ontological categories. The analogy the friend of PA appeals

to thus breaks down. In any case, to defend their favored alternative, the friend of PA has to tell a very complicated story. If so, it makes more sense to stick to the simpler story: when I believe that the brick in front of me is red, the object of my belief is just the red brick in front of me. Given this conception of the so-called propositional attitudes, Chisholm's version of the state-of-affairs view appears even less motivated.

But the state-of-affairs view about events need not be tied to PA. There are other ways to define states of affairs, and it is certainly possible to formulate a version of the state-of-affairs view that excludes the idiosyncratic features of Chisholm's version. Let's now examine this possibility. Kevin Mulligan and Fabrice Correia note that certain constructions in natural language seem to enable language users to refer to facts (2017, sec. 1.1). For example, by combining the phrase "it is a fact" with the clause "that Edmund is playing tennis", one seems to convey that the clause names a fact, i.e. Edmund's playing tennis. There is then the metaphysical question about the nature of facts. According to Mulligan and Correia, a major theory of facts takes a fact to be a state of affair that obtains. A state of affairs may or may not obtain; it does if "an object exemplifies a property or one or more objects stand in a relation" (2017, sec. 1.1 para. 6). In our example, the state of affairs of Edmund's playing tennis obtains because Edmund exemplifies the property of playing tennis. By fixing the referent of the term "state of affairs" this way, we can avoid committing ourselves to PA or any particular theory of propositional attitudes.

Mulligan and Correia list several metaphysical roles that a state of affairs might play, and I take the most important one to be the role of making propositions true (2017, sec. 1.4). States of affairs are structurally similar to propositions, but they must be sharply distinguished from propositions. While there is an entity that is the content of the sentence

“Edmund is playing tennis”, there is another entity that makes the content true. The former is a proposition, while the latter is a state of affairs that obtains. One might worry that this view trivializes states of affairs. It is tempting to claim, say, that the proposition  $\langle$ Annegret Kramp-Karrenbauer is a German $\rangle$  is made true by Kramp-Karrenbauer herself: the proposition is true because she is indeed a German. Wouldn't this claim render states of affairs metaphysically superfluous, if they are posited to fill the truth-making role? This worry may be lessened if we take into account the range of propositions that states of affairs can be considered to make true. According to Mark Textor, the possible outcomes of a statistical experiment are paradigmatic states of affairs (2016, sec. 1.2). If I draw a card from a deck of playing cards, one possible outcome is that I would draw the two of clubs. The outcome of my drawing the two of clubs is then a state of affairs. Since the probability distribution of a statistical experiment is an objective matter, this state of affairs exists whether or not I actually draw the two of clubs. Extrapolating from the claims made by Mulligan, Correia and Textor, we arrive at the conclusion that states of affairs also make propositions about probability true. Let  $X$  be an experiment in which I draw 13 cards from a deck of playing cards. Since the proposition  $\langle$ more than 10 cards drawn in  $X$  are clubs $\rangle$  corresponds to one of the possible outcomes, there is a state of affairs that can make the proposition true by obtaining. It is unclear that the truth-making role of this state of affairs can be played by any other entities. The suggestion that the cards themselves make the proposition true doesn't sound promising: if I draw 13 clubs, which of the clubs make the proposition true? It thus emerges that we can motivate the ontological commitment to states of affairs without accepting Chisholm's view. On the present view, if belief-contents are

propositions, then anyone believing Edmund to be playing tennis indeed counts as having a correct belief in virtue of the state of affairs. Despite this, the state of affairs should not be defined in terms of anyone's belief; it obtains even if it is somehow impossible for anyone to believe that Edmund is playing tennis.

Equipped with this alternative conception of states of affairs, we can now identify an event with the *obtaining* of a state of affairs. Edmund's current playing of tennis is an event, and it can be identified with the current obtaining of the state of affairs that Edmund is playing tennis. We then have a theory that unifies events, facts and states of affairs, and parsimony is a great virtue for any metaphysical theory. Call this theory the *unified theory* (UT). Why not accept UT? The problem, I submit, is that the "obtainings" of states of affairs are not individuated the right way. Their properties simply fail to correspond to those of events. To see this, consider a distinction made by Textor (2016), which is in turn drawn from Pollock (1984). A state of affairs is either *transient* or *nontransient*. A transient state of affairs is one that meets the following modal criterion: it is possible for there to be times at which the state of affairs obtains as well as times at which it doesn't. A state of affairs is nontransient if and only if it is not transient (Textor 2016, sec. 2.3 paras. 4-5). For example, the state of affairs of Edmund's playing tennis is transient, while Edmund's playing tennis on January 1, 2000 is nontransient. As we shall see, both kinds of states of affairs differ from events in important ways.

Let's begin with nontransient states of affairs. Consider again Edmund's playing tennis on January 1, 2000; call it *s*. Can UT identify the event that is Edmund's playing of tennis with the obtaining of *s*? As a nontransient state of affairs, *s* always obtains if it obtains at all. Identifying the event with the obtaining of *s* is therefore tantamount to saying that the



event goes on forever. This is obviously problematic. As a rejoinder, one could dispute the decision to think of  $s$  as a nontransient state of affairs. Why not say that  $s$  only obtains on January 1, 2000? We cannot do that if the main role played by states of affairs is that of making propositions true. If the proposition  $\langle \text{Edmund will play tennis on January 1, 2000} \rangle$  was ever true before January 1, 2000, then it was presumably made true by  $s$ . Similarly, if the proposition  $\langle \text{Edmund played tennis on January 1, 2000} \rangle$  continued to be true after January 1, 2000,  $s$  would be the best candidate for its truth-maker. Since a state of affairs has to obtain to make a proposition true, it follows that  $s$  always obtains. We are then led back to the conclusion that states of affairs of the kind under consideration are eternally-obtaining. Observations like this could easily be used to motivate the stronger view that states of affairs are non-spatiotemporal, insofar as one assumes that nothing spatiotemporal has the eternal truth-making features exemplified by states of affairs. Indeed, the stronger view was historically influential, as witnessed by Barry Smith's work on *Sachverhalte*, in which he interprets Adolf Reinach as taking states of affairs to be set-like abstract objects: "Like sets, *Sachverhalte* are built up (*inter alia*) out of ordinary objects in a way that somehow suspends the mutability of the latter" (B. Smith 1989, 64). If states of affairs are immutable, it is hard to see how their obtaining can be construed as the occurrence of events. After all, events are commonly regarded as the locus of change.

Does it mean that we should take an event to be the obtaining of a transient state of affairs? That won't do, either. Consider again experiment  $X$ , in which I draw 13 cards from a deck of playing cards. Let  $s_1$  and  $s_2$  be two transient states of affairs: the former is my drawing more than 10 clubs in  $X$  and the latter my drawing more than 12 clubs in  $X$ . Since

the probability of  $s_1$ 's obtaining is different from that of  $s_2$ 's obtaining, these two states of affairs are distinct. Suppose I carry out  $X$  at 10 am on January 1, 2020 and draw 13 clubs; it follows that  $s_1$  and  $s_2$  both obtain at 10 am on January 1, 2020. Given that  $s_1$  and  $s_2$  are distinct, the obtaining of  $s_1$  is distinct from the obtaining of  $s_2$ . According to UT, I have made two events happen by drawing 13 cards. But there is only one event. Even modal considerations wouldn't enable us to establish the existence of two distinct events: in any possible world in which I draw more than 12 clubs by drawing 13 clubs, that very drawing event is identical to the event of my drawing more than 10 clubs by drawing 13 clubs. This shows that UT gives the wrong identity condition of events. Of course, the friend of UT could weaken the theory: she could argue that the obtaining of a state of affairs is necessary, but insufficient, for the existence of an event. The friend of UT could then say that neither the obtaining of  $s_1$  nor the obtaining of  $s_2$  is sufficient for the existence of an event; only the obtaining of my drawing 13 clubs is. The problem is that even if  $s_1$  and  $s_2$  are necessary for the existence of the drawing event, they seem irrelevant to the nature of the drawing event. That is, they don't appear to play any role in the identity condition of the drawing event; the identity of the drawing event seems to be completely determined by the state of affairs of my drawing 13 clubs. To defend UT, therefore, one has to assume that the existence of an event depends on something that has nothing to do with its nature. There may be ways to motivate such a view, but I don't see much appeal of it. What then emerges is that neither Chisholm's view nor UT captures the central features of events. For this reason, I shall not pursue the state-of-affairs view any further in this chapter.

## 2.4 The Instantiation View

### 2.4.1 *Kim's Account*

According to the instantiation view of events, events occur when properties are instantiated. Kim is commonly credited with initiating this approach to events. In Kim's account, events are taken to be "exemplifications by substances of properties at a time" (1976, 160; emphasis omitted). For example, the exemplification of rotation by the London Eye now is an event. Kim's use of the term "substance" is very liberal; it includes both physical objects and the matter that constitutes them (1976, 159). If we take the paradigmatic examples of substances to be such entities as buildings, people and planets, then "continuants" would be a better term for what Kim calls "substances".<sup>10</sup> Kim also has a somewhat relaxed criterion of the properties constitutive of events: he doesn't think that change is necessary for events (1976, 159–60). Therefore, the induced anesthesia in a patient could be an event, because the patient has the property of being anesthetic at the relevant time.

What advantages might the instantiation view enjoy over the state-of-affairs view and the space-time view about events? I think the best way to motivate the view is to consider an example provided by Carol Cleland. Take a sphere that changes color while rotating; the rotation coincides spatiotemporally with the color change (Cleland 1991, 230). On Kim's view, there is a simple way to understand the case: it involves two events, which are the

---

<sup>10</sup> I will understand the term "continuant" the way Peter Simons defines it. Here's Simons' definition: "A *continuant* is any object which exists in time and which has no temporal parts, that is, parts which exist solely because of its existing at a certain time. Continuants persist by enduring" (2000a, 59).

sphere's exemplification of rotation at the relevant time and its exemplification of color change at the same time. The analyses offered by the other two views are more contrived. Given the spatiotemporal coincidence, the space-time view would have it that the rotation and the color change are the same event; this is a very undesirable verdict. The state-of-affairs view is faced with a different problem: in addition to the sphere and the properties, the view has to posit two entities called "the obtaining of the sphere's rotating" and "the obtaining of the sphere's changing color". The comparison suggests that the instantiation view has more intuitive appeal than the other two views.

This is not to say that Kim's version of the instantiation view is without problems. I don't think that every event is the instantiation of a property by a continuant. To begin, there are some putative counterexamples to this claim. As Casati and Varzi point out, it has been argued that weather events and the like do not involve any objects – though they also note that these may not be real counterexamples if we opt for a broader conception of objects (2008, 42). Given that Kim is willing to count quantities of matter among the continuants accommodated by his theory, he could argue that weather events such as storms still involve continuants – how could a storm possibly occur without some quantity of air?<sup>11</sup> This may be true, but it is irrelevant unless Kim makes the further claim that the storm is a property of the quantity of air. This further claim is problematic. Suppose a storm persists for two hours; let  $Q$  be the quantity of air that allegedly instantiates the storm. I claim that  $Q$  is individuated

---

<sup>11</sup> I will follow Evnine's use of the term "quantity". Evnine understands the term the way Helen Cartwright (1970) does: it refers to a collection of matter, not the amount of matter in such a collection (Evnine 2016, 3 fn. 5). Hence, "a quantity of sugar" doesn't denote anything like 5 tablespoons or 2 kilograms; it simply denotes a collection of sugar.

by the storm. For we could ask: what makes  $Q$  a single quantity of air? For every moment in the two-hour period, there is a quantity of air that occupies the same spatial region as the storm, and these quantities of air are all different from each other. What makes them a single quantity of matter? The most natural answer is: each of them occupies the same spatial region as the storm at some moment in time. This is what the air molecules in  $Q$  have in common. If so, we should say that the numerical identity of  $Q$  depends on that of the storm, not the other way around. But if the numerical identity of  $Q$  depends on that of the storm, there is little reason to think that the storm is a property of  $Q$ . The further claim Kim has to make is thus untenable.

To be sure, the friend of Kim's view could argue that I have confounded epistemological and metaphysical issues. For us to identify  $Q$  in our thoughts, we have to invoke the storm. But it doesn't follow that the numerical identity of  $Q$  depends on that of the storm. We could simply assume that any two quantities of matter constitute a single quantity of matter. This is Tyler Burge's view. He notes that stuff like water is *cumulative* (1977, 112). If we have two quantities of water, then their collection is also a quantity of water.<sup>12</sup> This makes stuff very different from paradigmatic individuals: the collection of two human beings is not another human being. But there are at least two ways to conceptualize a

---

<sup>12</sup> Burge's theory of stuff is part of his more general theory of aggregates. His official definition of being cumulative is given in the following *principle of cumulative reference*:

$$\varphi(x) \ \& \ \varphi(y) \ \rightarrow \ \varphi([z: zax \vee zay]) \quad (\text{Burge 1977, 112})$$

In Burge's formalism,  $x, y$  and  $z$  are aggregates,  $\varphi$  is any one-place predicate that describes stuff,  $a$  is the two-place predicate of being a member-component, and "[ ]" stands for totality. Therefore, the principle of cumulative reference reads: if  $x$  and  $y$  are stuff (with certain characteristics), then the totality of their member-components is also stuff (with the specified characteristics).

quantity of matter. To see this, consider Lake Michigan and River Rhine. On the one hand, we could think of all the water molecules in them as constituting a single quantity of water. On the other hand, we could take Lake Michigan and River Rhine to be two distinct waterbodies. The two waterbodies have very different properties – there are many ships that could sail on Lake Michigan but not on River Rhine. It would hence be problematic to say that they constitute a single waterbody. To resume our discussion of the storm case, I believe that the second way to conceptualize quantities of matter is the more pertinent one. For a storm to occur in an air mass, the mass must have certain properties. If the mass were regarded as a homogenous collection of air molecules, the occurrence of the storm in the mass would be inexplicable. So construed, however, a quantity of air doesn't just combine with any other quantity of air. It is simply implausible to hold that an air mass in Siberia and one in the Indian Ocean constitute a single air mass. If so, what enables us to speak of  $Q$  as *the* air mass in which the storm occurs? There is not a single air mass that stays unchanged during the entire storm. We can then run our earlier argument again and conclude that the storm is what makes the air masses a single air mass. My point thus stands: the numerical identity of  $Q$  depends on that of the storm, so to think that the storm is a property of  $Q$  is to have things backwards. Kim's attempt to identify an event with the instantiation of a property by a continuant fails.

#### 2.4.2 *Cleland's Account*

A sophisticated alternative to Kim's version of the instantiation view is defended by Cleland. Her account differs from Kim's in at least two respects. First, while Kim allows the possibility that some events are not accompanied by changes, Cleland holds that all events

are changes (1991, 231). Therefore, though Kim would regard my stay at a hotel as an event, Cleland would not. Second, while Kim takes events to be property exemplifications by physical continuants, Cleland takes events to property exemplifications by tropes (1991, 237–38). *Tropes* are particular instances of properties such as the temperature in my room and the whiteness of the London Eye. A useful way to conceptualize tropes is provided by Peter Simons, who takes them to be “a kind of dependent concrete particular” (1994, 557). Tropes are dependent because they don’t exist on their own,<sup>13</sup> and they are concrete in that they have spatiotemporal locations. A trope is therefore different from both the property of which it is an instance and the object on which it depends. On Cleland’s view, an event such as an onset of depression is not a change in a person but a change in the mental condition of a person, which is a trope.

To make sense of the idea that events are property exemplifications by tropes, Cleland assumes that properties come in different degrees of specificity. For example, rotation is a specific kind of motion. The standard way to characterize their relation is to say that rotation is a determinate of motion while motion is the determinable of rotation.<sup>14</sup> Following Stephen Yablo’s terminology, we can say that rotation *determines*, or bears the *determination relation* to, motion (1992, 252). We are now ready to examine the details of Cleland’s account. The centerpiece of her account is the notion of a *concrete change*. Let  $p$  be a

---

<sup>13</sup> The criterion of dependence probably has to be qualified. Simons himself accepts the existence of trope bundles that are not dependent on anything else (1994, 567–68). I see no harm in treating these trope bundles as complex tropes, but such complex tropes would then be capable of existing on their own. This shouldn’t be a serious problem, because we can still say that *simple* tropes are dependent particulars.

<sup>14</sup> The distinction between the determinate and the determinable is often attributed to W. E. Johnson (1921).

trope that is an instance of  $P$ . Suppose  $Q_1$  and  $Q_2$  are distinct properties that bear the determination relation to  $P$ . Then the successive exemplification of  $Q_1$  and  $Q_2$  by  $p$  is a concrete change, where  $Q_1$  and  $Q_2$  are called the *initial state* and *terminal state* of the change (1991, 238). For example, suppose that Germany's GDP will triple between 2040 and 2050: its GDP will grow to 30,000 billion U.S. dollars by the end of 2050 from its 2040 value of 10,000 billion dollars. To apply Cleland's definition to this case, note that GDP is a property of an economy. We can thus take GDP to be  $P$  and Germany's GDP to be  $p$ . What about  $Q_1$  and  $Q_2$ ? Consider the property of having a GDP of 10,000 billion dollars and that of having a GDP of 30,000 billion dollars. Since these two properties bear the determination relation to GDP, we can take them to be  $Q_1$  and  $Q_2$ . It follows that the tripling of Germany's GDP between 2040 and 2050 is a concrete change: it is a pair of property exemplifications in which Germany's GDP first exemplifies a value of 10,000 billion dollars and then a value of 30,000 billion dollars. Notably, changes like this are called "concrete" because they need to be sharply distinguished from types of change. While the tripling of GDP is a type of change, the tripling of Germany's GDP between 2040 and 2050 is a token of the type. This definition of a concrete change captures all the essentials of Cleland's account – she simply identifies events with concrete changes (1991, 245).

How does Cleland defend her account? One of the arguments she offers is that her account is able to accommodate those events that occur independently of physical continuants. As examples of such events, she cites the "purely temporal world of disembodied melodies, booms, bangs, shrieks, etc." in the thought experiments presented by P. F. Strawson (1959), as well as "that shriek, this flash, that desire" and "fluctuations in gravitational and electromagnetic fields" (1991, 230–31). I certainly agree with Cleland that



these examples put pressure on accounts like Kim's.<sup>15</sup> However, even if the existence of physical continuants is not necessary for the occurrence of events, it doesn't follow that events need to be conceived of as Cleland's concrete changes. Indeed, why should we even say that change is necessary for events? Cleland's answer is that this is how events are conceptualized in everyday life (1991, 231). Since a change-based theory of events better coheres with the common-sense worldview, it is to be preferred to theories that accommodates changeless events.

I think Cleland's claim about the everyday conception of events is false. Common sense has it that events are causes and effects, but causes and effects need not be changes. As Lewis puts it, "We cannot afford to count the unchanges as nonevents, for the unchanges may be needed to complete causal histories" (1986c, 261).<sup>16</sup> Suppose Carl gets bored after staying in his room for the entire day. Carl's stay in his room hardly counts as a change of any kind. However, it is the stay that causes Carl's boredom; without treating the stay as a cause, his boredom would become inexplicable. It follows that changeless events are still events. Cleland could respond by saying that the above considerations are not really part of the common sense; rather, they are philosophical attempts to make our everyday conception

---

<sup>15</sup> Though I agree with Cleland's claim that events need not involve physical continuants, I am not sure that Cleland has made a strong case for this claim. As we have seen, Kim construes substances broadly such that they include quantities of matter. It is therefore insufficient to just cite such phenomena as sounds and flashes; one needs to explain why these phenomena cannot be understood as property exemplifications by quantities of matter. Cleland hasn't done that. Among Cleland's putative counterexamples, the only one that is straightforwardly incompatible with Kim's account is the case of purely temporal worlds. However, I am not inclined to believe that these worlds are really possible. Even if I am wrong, the "sounds" in those worlds probably differ in kind from the sounds in the actual world. Why think that they have anything in common? For these reasons, I don't think Cleland's arguments are successful.

<sup>16</sup> A similar concern is raised by Kim (1976, 159–60).

of events more precise. This response isn't really defensible. There is no straightforward way to decide what counts as part of the common sense and what doesn't – philosophers' intuitions are by no means the most authoritative guide in our attempt to make such decisions. One of the more reliable guides is natural language, and, as mentioned above, the place of events in natural language can be clarified with Davidsonian semantics of verb predicates. But the semantic values of verb predicates need not involve changes. Witness the following inference:<sup>17</sup>

(Premise) Carl reluctantly stayed in his room.

$$\exists e (\text{STAY}(\text{Carl}, e) \wedge \text{RELUCTANT}(e) \wedge \text{IN}(e, \text{Carl's room}))$$

“There is some  $e$  such that  $e$  is a stay,  $e$  is reluctant, and  $e$  occurs in Carl's room”.

(Conclusion) Therefore, Carl stayed in his room.

$$\exists e (\text{STAY}(\text{Carl}, e) \wedge \text{IN}(e, \text{Carl's room}))$$

“There is some  $e$  such that  $e$  is a stay and  $e$  occurs in Carl's room”.

This is a valid argument, and treating Carl's staying as an event makes explicit why the argument is valid. If so, contrary to what Cleland says, change isn't a necessary component of the everyday conception of events.

If the only problem with Cleland's account is her requirement that events be changes, then we could just drop this requirement. We could retain the idea that an event is

---

<sup>17</sup> This, of course, is a much simplified representation of the inference. Tense is ignored, and no attention is paid to the syntax. Those complications can nevertheless be set aside for the present purposes.

defined by its initial and terminal states, each of which is a property exemplification. While Cleland insists that these two states must be different, we could allow the possibility that they are the same. On this revised account, my two-day stay at a hotel counts as an event: its initial state is my stay at the hotel on the first day and its terminal state is my stay on the second day. Though this revised account seems promising, it is faced with a difficulty raised by Paul Needham (2013). Needham argues against Cleland and maintains that events must be sharply distinguished from states; if his argument is sound, there is no hope for us to succeed in describing an event in terms of its initial and terminal states.<sup>18</sup> Needham defends his claim by appealing to thermodynamics. In what follows, I shall examine two of Needham's main arguments.

Here's Needham's first argument: if we understand an event as the transition from one state to another, then we fail to make sense of cyclic processes (2013, 401). I may heat a cup of water and then leave it to cool down until it returns to the initial temperature. This process is clearly different from one in which I do nothing to the water. Nevertheless, the two processes would have identical initial and terminal states. It follows that a process cannot be defined by the states at its endpoints, however close the two endpoints are in time. In his second argument, Needham notes that a process might not correspond one-to-one with the correlative state-change (2013, 404–5). One possible state-change is the increase in temperature by 20 degrees Fahrenheit, but this state-change does not define any actual process. How much thermal energy it would take to bring about such a change

---

<sup>18</sup> Needham's main concern is with what he calls "processes". However, what he means by the term doesn't seem different from what others mean by "events", and he often appears to use the terms "processes" and "events" interchangeably.

depends on the heat capacity of the heated object. Take two objects, one made of copper and the other gold. Even if the temperatures of both objects are raised from 40 to 60 degrees over the same period, their changes in fact consist in converting different amounts of energy. Moreover, any temperature change might result from mechanical work rather than thermal energy. Despite the absence of any hot objects, the temperature of a quantity of air may rise because of an increase in its internal pressure. Therefore, while one may use the concept of a 20-degree increase in temperature to refer to many processes, the referent of this concept is determined by a variety of contextual factors. The concept does not uniquely pick out any type or token of process. It is thus futile to conceptualize a process by using state concepts.

I am not convinced that Needham's arguments are sufficient to show that events cannot be individuated by states. To be sure, his first argument demonstrates the importance of temporal continuity: an event is individuated by the entire series of states it goes through, not just by the initial and terminal states. This is indeed an important insight, and it shows why a theory like Cleland's doesn't do justice to the temporal nature of events. Even so, Needham's point can be accommodated without drawing a sharp distinction between events and states. We can formulate what may be called a continuity requirement for event as follows. Define  $TL(\cdot)$  to be a function that maps an event to its temporal location. Then for any event  $e$  and its temporal parts  $e_1$  and  $e_2$ , if the mereological sum of  $TL(e_1)$  and  $TL(e_2)$  is not identical to  $TL(e)$ , then  $e$  is not individuated by  $e_1$  and  $e_2$ . When applied to Cleland's account, this requirement says essentially the same thing as Needham's point. Let  $e_1$  and  $e_2$  be the initial and terminal states of an event; the requirement says that the event is not individuated by them. The requirement thus captures Needham's insight. Nevertheless, it is

compatible with the possibility that some events consist of series of states. If so, though Needham's first argument establishes the continuity requirement, it doesn't undermine our revised account.

What about Needham's second argument? Its conclusion is that given any series of temperatures, continuous or not, the series is not identical to any particular event of heating. However, it does not follow that the series is not identical to any event whatsoever. It may very well be that a temporal part of a heating process is always a complex event, of which temperatures are only components. Needham would object that temperatures cannot even function as such components. In thermodynamics, states like temperature are undefined for a system that is not in equilibrium. Hence, while being heated, a body is not in any state. Its putative states are "more realistically treated as patterns of states of small parts of the body, continuously varying over space, which adjust as the process proceeds unhindered until a uniform state of equilibrium under the prevailing constraints is attained" (2013, 411). Since temperatures do not even exist in a heating process, they are certainly unable to function as components of any kind. However, Needham's own remark indicates a problem with the objection. Why cannot we define a complex state for the heated body in terms of the states of its parts? It may be that the complex state has no place in thermodynamics, but the state certainly has a place in metaphysics. If the parts of the body compose a complex entity (i.e. the body), it is reasonable to assume that the states those parts are in also compose a complex state. This complex state is a temporal part of the heating process. As it turns out, Needham's second argument again fails to establish that events cannot be individuated by states.

Where does this leave us? Should we simply take Cleland's account and drop the requirement that events be changes? This may be tempting, but further reflection shows that a simpler account is preferable. Consider a changeless event, such as an official's acquiescence in government corruption. Should we say that this event is an exemplification of a property by a trope? Maybe the relevant trope is an instance of acquiescence, and the trope instantiates the property of being directed towards government corruption. Or maybe the relevant trope is a mental state, and the mental state instantiates the property of being an attitude of acquiescence. These descriptions of the event seem very contrived to me. Wouldn't it be more apt to simply say that the event *is* a trope? Taking this approach, we could describe the event as an instance of acquiescence in government corruption, which is a trope located where the official is. If this approach is plausible, then Needham's claims are too strong. There are ways to define events in terms of states without falling prey to the problems plaguing Cleland's account. I think this is exactly what Bennett's account does, so I will now turn to Bennett's version of the instantiation view.

#### 2.4.3 *Bennett's Account*

Bennett differs from the other theorists we have considered so far in that he pays more attention to natural language semantics. An important point stressed by him is that there is no foolproof way to read off the tropes constitutive of an event from a name of the event. To state Bennett's view more precisely, we need to define a new term. Consider event *e* and property *P*. Bennett says that *P* constitutes *e* if and only if "the whole intrinsic truth about *e* is that it is an instance of *P*" (1988, 93; emphasis omitted). To avoid confusion with the constitution relation that is the central concern of hylomorphism about events, I shall

use the term *individuating properties* to denote properties that constitute events in Bennett's sense of "constitute". For an expression to name an event, Bennett argues, it often suffices for the expression to name some components of the individuating property of the event (1988, 93–94). For example, the expression "the cowbird's flap of wings" is a name of an event, even though the expression hardly specifies the entire cluster of properties constitutive of the event – a flap of wings is always done at a certain speed and in a certain direction, etc. These motion properties are also components of the individuating property of the cowbird's flap of wings, but they are not described at all by the expression "the cowbird's flap of wings". It then transpires that events do not have unique names; even if two expressions describe different properties, it does not follow that they name distinct events. This is why, given an appropriate context, "the cowbird's flap of wings" names the same event as "the cowbird's sudden movement". The moral is that we have to be careful when we attempt to determine the numerical identity of an event. Before we ask, say, whether my present experience and the present activity of my brain *are* the same event, we should first ask whether "my present experience" and "the present activity of my brain" *name* the same event. We risk confounding semantics and metaphysics if we fail to do so.

With Bennett's semantics in place, we can now discuss his metaphysics. Bennett takes an event to be a trope at a spatiotemporal region: "an event is the instantiation of a property *at a zone*. The zone will often be delimited by a substance and a time, but perhaps not always" (1988, 88). For example, if the property of pollution is instantiated throughout the spatial region occupied by a factory at time *t*, then the pollution in the factory at *t* is an event – this is the case on the assumption that the property of pollution is the kind of property whose instantiation is sufficient for the occurrence of an event. On Bennett's

account, there is no straightforward way to determine whether this assumption holds. And unlike Kim and Cleland, Bennett is unwilling to take a stand on the issue of whether events must always exemplify properties whose instances are accompanied by changes. His reason is given in his 1996 paper: our ordinary event concepts, as manifested in our patterns of inference, warrant neither a positive answer to this question nor a negative one (1996, 150–51). The imprecision of our event concepts also explains why, given any trope, there is no telling *a priori* whether the existence of the trope is sufficient for the occurrence of an event. All we can be sure about is that if something is an event, then it is a trope.

Does Bennett’s theory provide an adequate description of the nature of events? I do not think so. The greatest virtue of Bennett’s account is its simplicity, but it comes with a price. Given that Bennett cannot resort to Kim’s continuants or Cleland’s instances of determinable properties, it is unclear how he can account for the unity of a complex event. Consider the passage of a particular no-confidence motion in a parliamentary democracy. Suppose the event takes place over time interval  $T$  and consists of two phases: the prime minister addresses the members of the parliament and defends the government, and then the parliament members vote to pass the no-confidence motion in response to the prime minister’s speech. Suppose the motion cannot be legally passed if either phase is omitted. Call the event as a whole **no-confidence** and the two subevents **speech** and **vote**. Now we ask: what is the relation between **speech** and **vote**, on the one hand, and **no-confidence** on the other? It is my contention that Bennett’s account doesn’t have the resources to answer this question appropriately.

To see why, note that intimate relations obviously obtain among the three events. In any possible world where **speech** occurs but **vote** doesn’t, **no-confidence** fails to occur; the



same goes for the worlds where **vote** occurs but **speech** doesn't.<sup>19</sup> It thus seems that we cannot appropriately characterize the nature of **no-confidence** without invoking both **speech** and **vote**. The three events somehow form a whole. How should we understand this whole? The answer I prefer says that **speech** and **vote** jointly constitute **no-confidence** – **speech** and **vote** bear the matter relation to **no-confidence**. This is the analysis I shall develop in the next chapter. The constitution-based analysis is nevertheless unavailable to Bennett. He has indeed briefly considered the possibility that relations similar to material constitution may exist between events (1988, 124). However, as far as I know, he nowhere develops a theory based on such considerations. The best Bennett could do to accommodate the present observation is thus to defend the following hypotheses:

(The Fusion Hypothesis)

There is an event that is the fusion of **speech** and **vote**.

(The Identity Hypothesis)

The fusion of **speech** and **vote** is numerically identical to **no-confidence**.

The combination of the Fusion and Identity Hypotheses is in fact an attractive alternative to the constitution-based analysis: if Bennett's account has the resources to make this alternative work, maybe the constitution-based analysis is dispensable after all.

Unfortunately, this turns out not to be the case.

---

<sup>19</sup> In fact, I think it is more precise to say that **no-confidence** cannot occur without the occurrence of events *e* and *f*, where *e* and *f* fall under the kinds of which **speech** and **vote** are instances, respectively. For additional discussion, see section 3.5.2.

Let us set the Fusion Hypothesis aside for now and consider the Identity Hypothesis. We ask: on the assumption that **speech** and **vote** have a fusion (call the fusion **sum**), can Bennett identify their fusion with **no-confidence**? To answer this, we need to consider the relevant tropes. According to Bennett's semantics, we could refer to **no-confidence** by the definite description "the passage of the no-confidence motion", and it picks out a trope located in the parliament building. This trope is an exemplification of the event's individuating property. This property has many components, which not only includes the property of being the passage of a no-confidence motion but also those of pleasing the members of the opposition and being the last motion passed in the year, etc. Let NO-CONFIDENCE denote the individuating property of **no-confidence**. Similarly, let SPEECH, VOTE and SUM denote the individuating properties of **speech**, **vote** and **sum**. It is clear that SPEECH, VOTE and NO-CONFIDENCE are distinct properties; each of these properties can be instantiated independently of the others. Hence, whether or not the conjunction of SPEECH and VOTE has the same extension as NO-CONFIDENCE, the former property is not the same as the latter property. It follows that SUM is distinct from NO-CONFIDENCE, which means that the exemplifications of these two properties are different tropes. We can then apply Bennett's theory and conclude that **sum** and **no-confidence** are different events despite their spatiotemporal overlap. Bennett's theory fails to vindicate the Identity Hypothesis. Had he taken Kim's approach, he might be able to solve this problem by arguing that the complex property SUM is instantiated at the same time as NO-CONFIDENCE by the same unidentified substance. Whether or not this approach would turn out successful, it is not open to Bennett – substances do not appear in Bennett's theory.

Bennett could argue that my objection rests on the failure to appreciate what he has said about the semantics of event names. The definite description “the passage of the no-confidence motion” picks out an event that has many properties in addition to the property of being the passage of a no-confidence motion. If we examine these additional properties, it could turn out that SPEECH and VOTE are among them. If so, **sum** and **no-confidence** actually fail to name different events. It is unclear to me that this is really possible. Since SUM is formed by somehow combining SPEECH and VOTE, SUM doesn’t have any component that is not a component of SPEECH or VOTE. But the property of being the passage of a no-confidence motion is neither a component of SPEECH nor a component of VOTE (recall our assumption that the law requires the prime minister to defend the government before the parliament members make the final decision). It follows that the property is not a component of SUM, either. If so, “the passage of the no-confidence motion” cannot be a name of **sum**; the expression obviously names an event whose properties include that of being the passage of a no-confidence motion.

Perhaps there are ways to amend the semantics of event names and show that the ordinary-language names of **no-confidence** must also be names of **sum**. Such pursuits would be difficult in view of Bennett’s notion of an individuating property. Recall that SPEECH is not just the property of being a speech; it is the individuating property of **speech**, so its components may include such properties as being aggressive and being 10-minute long. It is hard to believe that the definite description “the passage of the no-confidence motion” has to secure its referent on the basis of such properties. However, even if one could come up with smart ways to defend the Identity Hypothesis for Bennett, Bennett

would not be able to defend the Fusion Hypothesis: his account doesn't have the resources to show that **speech** and **vote** have a fusion. Let's now turn to this issue.

As I see it, the most promising defense of the Fusion Hypothesis Bennett could give appeals to what he calls *zonal fusion*. A case of fusion is zonal if "the zone occupied by the whole is a fusion of smaller zones occupied by its parts" (1988, 143). Hence, in our example, two conditions must be met if **no-confidence** is to be regarded as the zonal fusion of **speech** and **vote**: (1) **no-confidence** is the fusion of **speech** and **vote**, and (2) the spatiotemporal region occupied by **no-confidence** is the fusion of the spatiotemporal regions occupied by **speech** and **vote**. What is the relation between these two conditions? Bennett offers no explicit answers. One possibility is that they are mutually independent. In that case, condition (1) requires us to identify a criterion of event fusion without appealing to their spatiotemporal locations. Bennett does suggest one such criterion, although he remains neutral on whether it should be accepted. According to it, if two events are parts of a larger event, there must be a causal chain running from one to the other (1988, 154).

Unfortunately, the criterion is not of much help in our example because **speech** can hardly be said to cause **vote**. As Lewis has famously argued, causal relations can be captured by counterfactual conditionals. In Lewis' account, the conditions that must be met for an event to cause another depend on whether the events in question occur in the actual world.

Consider events  $c$  and  $e$ . If they are non-actual, then  $e$  causally depends on  $c$  just in case that " $e$  would have occurred if  $c$  had occurred"; if they are actual, then  $e$  causally depends on  $c$  just

in case that “if  $e$  had not been,  $e$  never had existed” (Lewis 1986b, 167).<sup>20</sup> These conditions are not met by **speech** and **vote**. If **speech** had not occurred, **vote** could still occur; it’s just that **vote** would lose its legal significance. In addition, in a possible world where **speech** occurs, **vote** could still fail to occur – the prime minister could be giving the same speech at the same time for a different purpose. These observations suggest that **vote** does not causally depend on **speech**. If so, one cannot appeal to causation to defend the Fusion Hypothesis.

An alternative possibility is that conditions (1) and (2) are equivalent: **no-confidence** is the fusion of **speech** and **vote** if and only if the regions occupied by **speech** and **vote** compose the region occupied by **no-confidence**. Nonetheless, provided that unrestricted composition holds for spatiotemporal regions, this would imply that unrestricted composition also holds for events. And it does seem that any two spatiotemporal regions compose another: there is a spatiotemporal region that consists of Manhattan in 1990 and Manhattan in 2000. One might want to deny this because nothing seems to have a spatiotemporal location that coincides with this region; having such a location would require an entity to come into existence after ceasing to exist. Nevertheless, spatiotemporal regions are not individuated by the entities that occupy them. The O’Hare Airport and the Heathrow Airport were in the same space throughout 2018. There is thus a spatiotemporal region that contains both airports – the space in 2018. However, there is no one single thing whose spatiotemporal location coincides with this region. It follows that the unity of a spatiotemporal region does not depend on the existence of some entity whose

---

<sup>20</sup> Lewis distinguishes between *causation* and *causal dependence* (1986b, 167). Take three distinct events,  $e$ ,  $f$ , and  $g$ ; assume that  $g$  causally depends on  $f$  and  $f$  on  $e$ . In this case,  $g$  is caused by  $e$  whether or not  $g$  causally depends on  $e$ . This distinction doesn’t have much bearing on our present discussion.

spatiotemporal location coincides with the region. Hence, the aforementioned objection is misguided, and unrestricted composition does hold for spatiotemporal regions. On the assumption that conditions (1) and (2) are equivalent, the same holds for events.

This is an undesirable result. There are two reasons that unrestricted composition should *not* hold for events. First, there are numerous counterexamples to unrestricted composition, and Bennett himself has mentioned some. If there are two conference rooms in which scholars talk about issues in their field, do their activities constitute two distinct conferences or two parts of one conference? Bennett's answer is that it depends on the relations between these people – if all the scholars are characterized by “overlap of interest and concern” and “a relevantly shared causal history”, then their activities compose one single conference (1988, 124). If not, then there is nothing of which the activities in the two rooms are both parts. However, Bennett's answer is incompatible with unrestricted composition, which entails that the activities compose one single conference regardless of the relations between the scholars. If we accept Bennett's analysis, conferences constitute counterexamples to unrestricted composition. Second, unrestricted composition leads to problems for causal explanation. Suppose there are two conferences, one on German Idealism and the other on contemporary metaphysics. To explain why Francis, a devoted idealist, is upset after attending the second conference, we might appeal to the following observation: very few scholars in the conference take idealism seriously. The conference causes Francis' frustration because the conference has this property. Now, given unrestricted composition, the two conferences compose a complex conference, and the complex conference certainly does not have this property. It follows that the complex conference does not cause Francis' frustration. But I take the following to be a plausible assumption:

provided that event  $e_0$  is a part of event  $e$ , if  $e_0$  causes some event  $f$ , then  $e$  also causes  $f$ . If an inappropriate comment in a talk caused a controversy, then the talk certainly caused the controversy. Unrestricted event-composition contradicts this plausible assumption. For these two reasons, we should not accept unrestricted composition, which means that conditions (1) and (2) cannot be equivalent.

A third possibility is that the conditions are not equivalent but the satisfaction of condition (1) more or less depends on the satisfaction of condition (2). For example, it might be that **no-confidence** is the fusion of **speech** and **vote** if and only if the regions occupied by **speech** and **vote** compose a continuous region that is occupied by **no-confidence**. This criterion still posits too many events. Suppose **no-confidence** takes place right after the parliament passes the budget for the next fiscal year. The passage of the budget is spatiotemporally continuous with **no-confidence**, but there is little unity between the two events. Or consider some tourists' visit to the parliament building. The tourists may be taking pictures outside the parliament building when the prime minister address the parliament, but such spatial continuity does not make the tourists' activities and **no-confidence** a single event. If so, whatever the logical relations between conditions (1) and (2) are, spatiotemporal continuity is insufficient to bridge the conditions. To hold onto the third possibility, we need additional constraints, but none seems forthcoming.

As it turns out, though the case under discussion certainly involves the fusion of regions, Bennett's theory of zonal fusion cannot establish the numerical identity between **no-confidence** and the fusion of **speech** and **vote**. It then appears that Bennett's account of events does not have enough theoretical resources to defend the Fusion Hypothesis. Therefore, the simplicity that is the strength of Bennett's account is at the same time its

weakness. His account dispenses with the unnecessary entities posited in Kim's and Cleland's accounts, but it becomes inexplicable how the various elements of a complex event are unified the way they are.

## **2.5 Conclusion**

Despite the shortcomings of Bennett's account, I take it to be the most successful non-hylomorphic theory of events. Bennett's insight that tropes are constitutive of events is still promising; what is missing in his account is the tool to handle the unity of a complex event. Hylomorphism may provide such a tool. The best way to proceed, I suggest, is to develop a theory of events by combining Bennett's insight with some form of hylomorphism. This is what I will attempt in the next chapter.



## CHAPTER 3

### EVENTS: A HYLOMORPHIC PROPOSAL

In formulating the existence and identity conditions of events, one has to decide whether to invoke the relation of constitution. If one does, then one is committed to hylomorphism about events. This chapter defends a version of hylomorphism. On my view, just as a fresco may be constituted by a quantity of lime plaster, a court trial may be constituted by a series of announcements, arguments, questions and responses. To appreciate the similarity between the two cases, note that the various events just mentioned can be said to supply what is needed for the trial to proceed. Furthermore, these events enable the trial to go on without turning the trial into something reducible to the collection of these events: the trial could still have happened had some questions not been raised, and the various events could all appear in a hearing without turning the hearing into a court trial (as opposed to, say, a congressional hearing). The relation between the trial and the collection of the events is thus analogous to that between a fresco and a quantity of lime plaster – the latter supplies what is needed for the former to exist without making the former reducible to the latter. A theoretically fruitful way to exploit the similarity is to take the court trial to be literally constituted by the series of events occurring in it. The events are the matter of the trial; to put it differently, the series of events has the form of a trial.

The attempt to apply the notion of constitution to the metaphysics of events dates back to at least as early as Fine (1982), but it seems that Fine's paper has received relatively little attention in the literature on events. Recent years have seen some renewed interest in the constitution of events. Different views about event constitution have been advanced by

Crowther (2011), Ewnine (2016), Johnston (2006) and Jones (2013). What distinguishes my version of hylomorphism about events from the alternatives is my attempt to define the matter relation in terms of dependence relations. I shall argue that doing so resolves the difficulties faced by the other hylomorphic theories of events.

Here's the plan of the chapter. In Section 3.1, I will say a few words about the very idea of constitution and why it can be helpful to apply this idea to the analysis of events. Sections 3.2–3.4 examine the theories advanced by Crowther, Ewnine and Jones in detail. My own proposal is laid out in Section 3.5. After that, I respond to some potential objections in Section 3.6 before I conclude the chapter in Section 3.7.

### 3.1 Constitution in the Domain of Events

As we have seen, an entity bears the constitution relation to another if the former is the matter of the latter. But what is it for something to be the matter of another? Intuitively, such a relation can be characterized as “the relation between something and what it is made of” (Ewnine 2011, 212). When constitution is understood as *material constitution*, i.e. the kind of constitution relation that holds between matter and physical objects, its reality is hardly deniable: there is an obvious sense in which glaciers are made of ice, for example. To be sure, there may be the temptation to reduce the constitution relation to something else, say numerical identity; one may be inclined to say that a glacier is nothing over and above the quantity of ice that the glacier is made of. However, there are also strong reasons to resist this temptation. As noted by Lynne Rudder Baker, a physical entity and its underlying matter have very different persistence conditions (2007, 42). The glacier could survive even if one thirds of it were to be melted, but that would certainly destroy the quantity of ice. If so, the

relation between the glacier and the ice is not numerical identity. We are thus justified, at least in a *prima facie* way, in keeping the relation of material constitution in our ontology.

It is less clear how the notion of constitution is to be applied to events. Are events made of anything? What does this question even mean? To answer these, we have to consider what motivates an analysis of events in terms of constitution. Kit Fine has argued that such an analysis is useful for clarifying the relation between actions and bodily movements (1982, 102). As an example, he considers what events happen when one carries out the action of raising one's arm. The action certainly happens, and so does the bodily movement in which one's arm goes up. However, the bodily movement could have occurred by reflex; in that case, there would have been no intention, thus no action on one's part. Fine concludes that though the action and the bodily movement coincide spatiotemporally in the actual world, there are possible worlds in which only one of them exists. It is hence possible for two events to occupy the same spatiotemporal location while possessing different modal properties. As we have seen, this is precisely the reason that a constitution-based analysis is called for in the case of physical objects. If correct, then, Fine's view shows that the phenomenon of constitution can be found outside the realm of physical objects. It should be thought of as the unity between spatiotemporally coinciding but modally differentiated entities; as Baker puts it, constitution is to be understood as a kind of "unity without identity" (2007, 32). Since unity need not be construed exclusively in terms of physical objects, there shouldn't be any conceptual difficulties inherent in the idea of event constitution, even if the idea may, admittedly, strike one as somewhat surprising.

I think hylomorphism about events is very promising. If not for any other reason, hylomorphism directs our attention to the modal features of events. So much ink has been

spilled on issues about the modal features of physical objects, but it is far from obvious that we can successfully settle those issues without getting clear on the corresponding issues in the domain of events. Modal considerations are also important for the purpose of theorizing about events themselves: issues about identity are rarely separable from issues about modality. Therefore, it would seem that hylomorphism is in a unique position to deal with the tasks we set for ourselves in the preceding chapter: to look for the existence and identity conditions that events are subject to. Formulating a plausible account of constitution is nonetheless no easy task. In what follows, I will evaluate several existing proposals.

### 3.2 Crowther's Account

The first hylomorphic theory of events we shall consider is that offered by Thomas Crowther (2011). To evaluate Crowther's theory, we need to introduce a widely accepted distinction made by Zeno Vendler (1957). Vendler divides the verb phrases that can be combined with the progressive aspect into two groups. *Accomplishment terms* describe events that "proceed toward a terminus which is logically necessary to their being what they are" (1957, 146), while *activity terms* describe events that do not do so. For example, "sail into harbor" is an accomplishment term. If a boat is to sail into harbor, then it has to end up within a harbor – the arrival at a harbor is the terminus necessary for the occurrence of this event. In contrast, "sail on the river" is an activity term. From the claim that a boat sails on the river, it doesn't follow that there is anything in particular the boat has to do when the event reaches its end.

Vendler's distinction is widely accepted because it explains a variety of linguistic phenomena. There are many linguistic tests commonly used to separate accomplishment

terms from activity terms, and I shall mention two here. The first of them is offered by Vendler himself: the questions in which the two kinds of terms can appear are different (Vendler 1957, 145). Let  $S$  be a string of words such that any verb in  $S$  is uninflected. Only accomplishment terms can appropriately occur in questions of the form “how long did it take for  $S$ ?” For instance, question (1) below sounds fine, but question (2) does not:

- (1) How long did it take for the boat to sail into harbor?
- (2) How long did it take for the boat to sail on the river?

In contrast, only activity terms can be intelligibly used in questions of the form “for how long did  $S$ ?” As an example, question (3) below is natural, but question (4) is not:

- (3) For how long did the boat sail on the river?
- (4) For how long did the boat sail into harbor?

The test shows that the distinction between accomplishment and activity terms has semantic significance.

The second test appeals to what linguists call the *subinterval property*.<sup>21</sup> According to Lucas Champollion, “A predicate  $P$  has the subinterval property iff whenever it holds at an interval, it also holds at every one of its subintervals” (2017, 102).<sup>22</sup> The validity of the inference from (5) to (6) suggests that activity terms have the subinterval property:

---

<sup>21</sup> The idea of the subinterval property is typically attributed to Michael Bennett and Barbara H. Partee ([1978] 2004).

<sup>22</sup> A version of the test can be found in Vendler’s paper (1957, 145–46), but Champollion’s formulation is more precise.

- (5) Between 1 p.m. and 2 p.m., the boat sailed on the river.
- (6) At any time between 1 p.m. and 2 p.m., the boat sailed on the river.

But this is not the case for accomplishment terms. Consider the inference from (7) to (8), which is invalid:

- (7) Between 1 p.m. and 2 p.m., the boat sailed into harbor.
- (8) At any time between 1 p.m. and 2 p.m., the boat sailed into harbor.

This suggests that accomplishment terms do not have the subinterval property.

It has become standard to use the term “process” as another name of Vendler’s “activity”.<sup>23</sup> As we have seen, Vendler’s main concern is how the semantic properties of accomplishment terms differ from those of process terms. In contrast, Crowther’s goal is to defend the metaphysical distinction between the events described by the two kinds of terms, i.e. accomplishments and processes themselves. Such a defense is needed because the appeal to a terminus isn’t sufficient for showing that the events really fall under different kinds. Crowther uses the example of walking to make his point. Walking is a process, but it has a terminus: the stop of walking. After walking for a certain amount of time, one has to stop. Therefore, everything that happens in the walking process can be regarded as what one does to reach the point at which one can stop walking. It follows that the stop of walking is the terminus of a process of walking (Crowther 2011, 9–10). If so, processes have a terminus no less than accomplishments do.

---

<sup>23</sup> See Kearns (2011, 157).

To find a better way to differentiate accomplishments from processes, Crowther builds a theory of events that is hylomorphic in character. To do so, Crowther first appeals to linguistic observations to motivate a distinction orthogonal to the accomplishment-process distinction. On the one hand, terms like “walking” refer to temporally extended entities that “are like masses or substance-stuffs in not being count-quantifiable, only mass quantifiable. There can no more be two walkings or three runnings than there can be two golds or more than one rain” (2011, 16). Crowther calls the referents of such terms *temporal stuff*. His official definition of temporal stuff is “[w]hat things are doing throughout a period of time” (2011, 16). I take this to mean that kinds of temporal stuff are ways of taking up time. Since walking and running are two different ways of taking up time, they are different kinds of temporal stuff. On the other hand, terms like “walk” refer to temporally extended entities that “are countable, and cannot be mass-quantified. There can be more than one walk or more than one walk to the shops” (2011, 20). Crowther calls the referents of such terms *temporal particulars* and identifies events with them. Hence, both accomplishments and processes are temporal particulars.

What is the relation between temporal stuff and temporal particulars? Crowther identifies a temporal particular with a *completed* quantity of temporal stuff (2011, 19). For example, a sail is a completed quantity of sailing. To explain what this implies, we need to define the notion of a temporal boundary. A temporal boundary exists at a time if a quantity of temporal stuff appears or disappears at that time (2011, 20). For instance, suppose an airborne plane starts descending and then stops doing so at time  $t$ . Since the plane’s descending is a quantity of temporal stuff and the stuff disappears at  $t$ , a temporal boundary exists at  $t$ . We can now describe the idea of completion. Let  $x$  and  $y$  be temporal entities;

then  $x$  is completed by  $y$  if  $y$  can be a temporal boundary of  $x$  without making  $x$  “improper or defective” (2011, 27). Crowther illustrates the idea by appealing to an analogy between events and physical objects. If we, say, arbitrarily draw a line on a painting and cut the painting along the line, we get two defective paintings. It follows that paintings are not completed by arbitrarily drawn lines. As Crowther points out, similar scenarios can be found in the realm of events. A ballet performance is a quantity of ballet dancing. But if one arbitrarily divides a ballet performance into two, the resulting quantities of ballet dancing are at best incomplete performances. To avoid treating such incomplete performances as temporal particulars, Crowther defines temporal particulars to be those quantities of temporal stuff that are completed.

Crowther argues that the distinction between accomplishments and processes is akin to the distinction between physical objects and mere aggregates of matter (2011, 23). While dividing a painting into two leaves us two defective paintings, the situation is different if we divide a heap of sand into two. There is nothing defective about the resulting heaps of sand. It follows that mere aggregates of matter can be completed by any kind of spatial separation. According to Crowther, processes are the temporal analogues of mere aggregates of matter. In our earlier example, while the descent of the plane may be terminated either in midair or on the ground, these two options result in temporal boundaries of different kinds. If the pilot intends the plane to undergo a *process* of descending, then both options are fine. The process can be completed by either kind of temporal boundary. However, if the pilot intends the descending event to be an *accomplishment* of descending to an airport, then her only option is to terminate the descent on the ground. It thus emerges that while a process can be completed by any kind of temporal boundary, an accomplishment can only be completed by



boundaries of the kinds required by its nature. In Crowther's words, accomplishments "possess a temporal form provided by a temporal sortal or temporal principle of individuation for complete events of that kind" (2011, 34).

The extent to which Crowther engages with recent work in linguistics is impressive. However, his theory is not without problems. To begin, Crowther's theory has salient hylomorphic features, but it is unclear how he thinks about the form-matter relation. In his theory, all events have both form and matter. Processes are required to have the form of boundedness, as Crowther doesn't count a dispersed quantity of temporal stuff as a process: "A walk around, or a stretch of walking around is not stuff, but a bounded quantity of it" (2011, 25 fn. 39). The forms of accomplishments are even more complicated: such forms are as numerous and varied as what Crowther calls temporal sortals. But what exactly are temporal sortals? Take the descent of the plane. When the plane starts descending, is there any temporal sortal that is a component of the unfolding event? It seems that Crowther has to answer in the negative. Depending upon what happens in the future, the event could end up being an accomplishment or a process. Only then will the event become classifiable into a particular kind and feature a structure conforming to a temporal sortal. So construed, temporal sortals are properties exemplified by completed events; they do not exist in unfolding events.

This appears to conflict with what Crowther says about temporal stuff. Take walking and walking to a shop; he seems to think that they are different kinds of temporal stuff (2011, 20, 24). Therefore, before quantities of walking and walking to a shop are completed,

they already exemplify different temporal sortals.<sup>24</sup> This is inconsistent with my analysis in the preceding paragraph. To defend his position, Crowther has to claim that my analysis is wrong. But how does Crowther justify that? If I am walking to a shop, what I am doing doesn't seem different from what I would be doing if I were to simply take a walk. It follows that they should be regarded as the same kind of temporal stuff, because one cannot create two kinds of temporal stuff by carrying out one single kind of action. There is thus no reason to think that two different temporal sortals are exemplified here. Crowther has a response to this worry: "what is going on at a time... can be determined by what goes on over a larger period of time" (2011, 12 fn. 24). This, unfortunately, fails to address the worry. Crowther's claim only shows that a quantity of walking could either end up being a casual walk without a particular destination or a walk to a shop, not that a quantity of walking may have been a quantity of walking to a shop. Seeing things my way also accords with common sense better: if I am walking while debating with myself whether I should go to a shop, does such walking count as a kind of temporal stuff different from both walking and walking to a shop? The most reasonable answer is "no". Walking, walking to a shop and walking while debating with oneself whether to go to a shop are the same kind of temporal stuff; otherwise our ontology would include unnecessarily many kinds of temporal stuff. If so, my analysis should be upheld: temporal sortals are only exemplified by completed events. What Crowther says about temporal stuff makes his theory inconsistent.

---

<sup>24</sup> Or Crowther might say that quantities of walking to a shop already exemplify temporal sortals but quantities of walking do not. Either way, my objection stands.

My second objection has to do with Crowther's conception of constitution. According to him, all the properties actually exemplified by an event are necessarily exemplified by the event (2011, 36). This view has ramifications for his view on constitution. The constitution relation is not invoked to explain the relation between two spatiotemporally coinciding but modally distinguished events; there are no such events. Rather, it is invoked to describe the way temporal particulars exist in time. Drawing an analogy with particulars in space again, Crowther notes that having a spatial boundary is not a sufficient condition for being a particular in space. A hole has a spatial boundary, but it is not a particular the way a physical object is. To properly describe the spatial nature of a particular, one has to go beyond specifying its spatial boundary (2011, 37). On the one hand, one has to describe the matter that the particular is made of, because it is in virtue of its matter that the particular is spatially extended. On the other hand, one has to describe the way the matter of the particular is structured, because the particular has the spatial properties it does only on the basis of its spatial configuration. The same goes for temporal particulars. To properly describe the temporal nature of an event, one has to describe its matter and temporal structure. It is for this reason that we should speak of the constitution of an event.

As I see it, this conception of constitution makes the constitution relation redundant. What Crowther takes to be the matter and structure of an event need not be so described. They are just what goes on when an event unfolds and the kind an event falls under. These can simply be treated as properties of an event, and it is unclear how these properties are

different from the other properties an event might have, such as its participants,<sup>25</sup> terminus, temporal parts and spatial location. If those do not motivate an analysis of events in terms of constitution, neither do what Crowther takes to be the matter and structure of events. The issue of redundancy is aggravated by Crowther's stance on the modal nature of events. If an actual walk to a shop could not have been a casual walk without a particular destination, why should we posit such an entity as a quantity of walking that is different from the walk to a shop? Isn't such an entity needed only if we affirm the existence of something that actually constitutes a walk to a shop but could have constituted a casual walk? If we deny such a possibility, it make more sense to simply say that the actual walk to a shop has the same temporal parts as a possible casual walk in their initial phases. It is rarely denied that events have temporal parts. If it is enough to invoke temporal parts, there is no reason to be committed to the controversial idea of event constitution.

The friend of Crowther's view could argue that one need not resort to the modal considerations to motivate Crowther's conception of constitution. Instead, one could argue that the conception goes hand in hand with the distinction between temporal stuff and temporal particulars. This brings us to my next objection: I think the distinction is not sufficiently motivated, either. Crowther defends the distinction by noting that count nouns and mass nouns admit of different modes of quantification, but these considerations may have to do with a feature of the English language rather than a feature of reality. Metaphysics is not English syntax; English is just one language among the huge number of languages

---

<sup>25</sup> I follow Casati and Varzi in using the term "participant" to refer to any physical entity constitutive of an event (Casati and Varzi 2008, 37).

spoken in the world. Crosslinguistic data are desirable when one claims that the syntax of natural language mirrors reality. In Mandarin Chinese, nothing straightforwardly corresponds to the count nouns in English. One can draw a distinction between talking and particular talks in English by inflecting the word “talk”, but inflections do not exist in Mandarin. As Feng-hsi Liu points out, “syntactically Chinese nouns are similar to mass nouns in English, as they cannot combine directly with numerals, but must combine with classifiers” (2014, 154). The sentence in (9) is ungrammatical:

(9) 我 今天 做了 一 演講。

wo jing tian zuo le yi yian jiang.

I today did a talk.

“I gave a talk today.”

The reason is that the classifier *chian* is missing. When it is inserted between “a” and “talk”, the sentence becomes grammatical, as in (10):

(10) 我 今天 做了 一場 演講。

wo jing tian zuo le yi chiang yian jiang.

I today did an occasion of talk.

“I gave a talk today.”

The examples show that the Mandarin expression of “a talk (*yi chiang yian jiang*)” resembles “a glass of water” or “a gust of anger”; there is no way to grammatically omit the classifier. It is true that one cannot speak of “talkings” or “walkings” in Mandarin, but one cannot speak of “talks” or “walks”, either. Crowther’s linguistic argument for separating temporal stuff from

temporal particulars fails. To be sure, even if a distinction is not drawn syntactically, it doesn't follow that the distinction is not drawn semantically. And even if a semantic distinction is not drawn in ordinary contexts, it doesn't follow that the distinction cannot be drawn in technical contexts. My worry is just that Crowther appeals to a point about the English language to make a metaphysical claim. Doing so is fine insofar as there are no counterexamples from other languages or metaphysical theories, but in this case there are at least *prima facie* counterexamples. Crowther's argument is therefore too quick.

My fourth objection is that Crowther fails to address what I call the *problem of constitution by a plurality* (henceforth "the CP problem"): in virtue of what can multiple events jointly constitute a single event? Crowther himself allows the possibility of constitution by a plurality: "a temporal particular can be made of different stuff at different phases of its existence; the very same birthday party may have been composed of the exchange of small talk at  $t_1 - t_2$ , and debauchery at  $t_8 - t_{10}$ " (2011, 38). But how is that possible on Crowther's account, exactly? The small talk ends when the debauchery begins; call the ending of the small talk  $e$ . Since  $e$  amounts to a temporal boundary, Crowther's account entails that one of the following three states of affairs must obtain. First, there is a completed process of small talk whose temporal boundary is  $e$ . If so, the debauchery is just another event; there is no reason to hold that the birthday party is jointly constituted by the small talk and the debauchery. Second, there is a completed accomplishment constituted by the process of small talk; its temporal boundary is also  $e$ . Here the problem we encountered in the first scenario arises again. Moreover, it is unclear what the temporal sortal governing the accomplishment is. Third, there is an incomplete accomplishment that will only become completed if something other than small talk occurs. What is the temporal sortal here? The

only relevant one seems to be the sortal governing birthday parties. But why cannot there be a birthday party that involves nothing but small talk? Since none of the three possibilities entails that the small talk is a quantity of temporal stuff that remains to be completed by the debauchery, Crowther cannot explain the unity between the two. As we shall see, this problem is also faced by the other existing versions ofhylomorphism about events. It is indeed one of the main reasons that an alternative hylomorphic theory is needed.

### 3.3 Evnine's Account

Evnine's hylomorphic theory of events is not as general as Crowther's; it focuses exclusively on actions. Despite the narrower focus, however, the theory makes innovative use of the Aristotelian idea that formal and efficient causes are closely intertwined. Evnine tries to show that "actions are artifacts" (2016, 209). The central idea behind this claim, roughly put, is that actions depend on intentions the way artifacts do. Artifacts do not come into existence on their own; they need to be intentionally made. The existence and identity of an artifact thus depend on the intention to make the artifact. A monkey can smear materials of different colors on wet plaster, but what is created would not be a fresco – a fresco cannot exist if no one intends the end product of the smearing process to be a fresco. Actions feature a similar kind of dependence on intentions. Suppose I had never heard of laptops with a touchscreen. Even if I touch the screen of the laptop and then spread my fingers apart, what I am doing wouldn't count as an action of enlarging the on-screen image. Without the intention to enlarge the image, actions of enlarging the image cannot occur. These suggest that the relation between an action and the intention of its agent is akin to that

between an artifact and the intention of its maker. If we can plausibly speak of an artifact as made by its maker, it is no less plausible to speak of an action as “made” by its agent.

What does one have to do to make something? A natural answer is that one has to work on some matter and give it a particular form. But what would be the matter and form of an action? A possible answer is sketched by Fine. As we have seen, he takes one’s action to raise one’s arm to be distinct from the bodily movement co-occurring with the action. The two events are nevertheless entwined: the action occurs if the bodily movement exemplifies the property of being done with the intention to raise one’s arm. The action can therefore be viewed as a hylomorphic entity dependent upon the bodily movement and the property – while the former is the matter of the action, the latter is its form (Fine 1982, 102). Evnine is unwilling to conceptualize actions this way.<sup>26</sup> He maintains that actions have matter but no forms. Therefore, though an action is different from the co-occurring bodily movement, their difference is not a difference in form but one in the means of production. While an action of raising one’s arm can only be produced with the intention to raise one’s arm, a bodily movement in which one’s arm is raised need not be produced with the intention to have one’s arm raised. Evnine calls his own theory *amorphic hylomorphism*, according to which hylomorphic entities “fall essentially under certain kinds and must be understood in terms of the kind-related processes of work on their (original) matter by which they come to exist” (2016, 12). Such a treatment of actions is again motivated by the

---

<sup>26</sup> As far as I can tell, Evnine doesn’t explicitly argue against Fine’s hylomorphic analysis of actions. Therefore, Evnine’s main dissatisfaction with Fine’s analysis seems to be the framework in which Fine’s analysis is put forward, i.e. the theory of what Fine calls “qua objects”. See Evnine (2016, chap. 2) for his discussion of Fine’s qua objects.



analogy with artifacts. The main difference between, say, pasta and bread is arguably that they are produced in different ways. In Evidentiary's view, the emphasis on the process of making makes his account more Aristotelian than many forms of contemporary hylomorphism; his theory respects the Aristotelian view that what something is is inseparable from how it comes to be.

There are two ways in which an agent "makes" an action, depending on whether the action is basic. Evidentiary takes a basic action to be an action that one can perform without first performing another action. Suppose Anton winks at a colleague to show friendliness. In this case, the action of winking functions as a means to perform the action of showing friendliness. The action of showing friendliness is thus non-basic, while the action of winking is basic. Evidentiary claims that non-basic actions are made of other actions. Let  $F$  name a non-basic action and  $G$  an action that one performs for the purpose of performing  $F$ . Then to make  $F$  is to perform  $G$  with the twofold intention to perform  $F$ , on the one hand, and to perform  $F$  by performing  $G$ , on the other (2016, 232). In the example just mentioned, Anton succeeds in making the action of showing friendliness because he winks at his colleague with the intention to show friendliness, on the one hand, and he intends to show friendliness by means of winking, on the other. The action of winking thereby bears the matter relation to the action of showing friendliness; the former constitutes the latter. What it takes to make a basic action is somewhat different. To make a basic action  $H$ , an agent has to perform  $H$  intentionally, and her intention must cause appropriate bodily movements that enable her to perform  $H$  (2016, 233). Insofar as Anton's intention to wink causes his winking, Anton succeeds in making the basic action of winking. It is unnecessary for Anton to have the intention to wink in any particular way; he need not intend to wink by means of winking at

the rate of ten times per minute, for example. When it comes to basic actions, therefore, the relations of the matter relation are not actions: basic actions are constituted by bodily movements.

Intentions figure prominently not only in the existence conditions of actions but also in their identity conditions. Take possible actions  $A_1$  and  $A_2$ ; suppose they are performed with intentions  $I_1$  and  $I_2$ , respectively. Then  $A_1$  and  $A_2$  are numerically identical if (a)  $I_1$  and  $I_2$  are the same intention, and (b)  $A_1$  counts as a kind of action that fulfills  $I_1$ ; similarly for  $A_2$  and  $I_2$  (2016, 245). For example, consider possible worlds  $w_1$  and  $w_2$ . At  $w_1$ , Anton winks at a rate of 10 times per minute; at  $w_2$ , he winks at a rate of 15 times per minute. Insofar as the winking actions are performed with the same intention, they are numerically identical. Now suppose that because of some neurological disorder, the same intention causes Anton to open his mouth at  $w_3$ . Since the action of opening his mouth by no means fulfills his intention, the action is numerically distinct from those performed at  $w_1$  and  $w_2$ . Intentions also function to unify multistep actions. Evnine thus has a response to what I call the CP problem: “What accounts for the existence of a single action of which the other actions are the matter is the overarching intention that guides the performance and sequence of the acts that are the matter” (2016, 221). If Edmund intends to greet a friend by first saying hello and then giving the friend a hug, his intention unifies the actions of saying hello and hugging, turning them into an extended action.

I have two main concerns with Evnine’s account. First, while the analogy between artifacts and actions is suggestive, the analogy may not be strong enough to warrant the application of his amorphic hylomorphism to actions. Second, Evnine takes successively unfolding actions to be unified by an overarching intention, but it is unclear that intentions

can play such a unifying role. I shall spell out the worries in turn. First, Evnine contends that his theory of artifacts also accounts for actions. He attempts to show this by exploiting the analogy between artifacts and actions. If we set aside Evnine's treatment of physical artifacts, however, amorphic hylomorphism about events is a surprising position, to say the least. If one claims that an action may bear the matter relation to another, one is immediately faced with the question: what is present in the constituted action but missing from the constituting action? The simplest answer is "a form", however forms are to be construed. However, as we have seen, Evnine chooses a different answer: nothing is missing from the constituting action. I think there are reasons to prefer the first answer to Evnine's. For one thing, one could then follow the lead of Fine and identify the forms of events with properties. Given that properties have been extensively studied in metaphysics, one can defend hylomorphism with existing resources. For another, it is hard to see why we shouldn't reify kinds. Amorphic hylomorphism entails that two actions cannot be of the same kind if they are performed with different kinds of intentions. It follows that for every action  $x$ , there is kind to which  $x$  belongs but the matter of  $x$  does not. What prevents us from identifying that kind, or the property of being that kind, with the form possessed by  $x$  but not by its matter? I am not sure how Evnine would respond. Admittedly these considerations do not refute amorphic hylomorphism. Nevertheless, they at least show that some version of Fine-style hylomorphism is to be preferred unless amorphic hylomorphism affords a superior analysis of the matter relation.

Unfortunately, I don't think it does. The conception of matter Evnine has been working with when he discusses the matter of actions isn't sufficiently motivated. As Evnine himself points out, there is a reason that one may be unwilling to speak of bodily movements

as a kind of matter on which one can work (2016, 229–30). When Carl makes, say, an hourglass, the matter Carl uses consists of entities that are not brought into existence by Carl himself but can be manipulated by Carl. The sand and wood are simply found in nature. When Carl makes an action, however, the matter doesn't consist of anything readily available for Carl to manipulate. Carl certainly manipulates his body, but that is not the matter of any action. The matter is Carl's bodily movements, which are supposed to be caused by Carl's intentions. If so, the following analysis would be more compelling than Eynine's: Carl brings two events into being instead of making the one on the basis of the other. Eynine has a response to this worry: the difference between the matter of actions and that of physical artifacts only shows that there are different kinds of matter. When making physical artifacts, "we must work with what is already there" (2016, 233); the matter we manipulate in this case is given to us. When making actions, however, we can create the kind of matter we want to use. We don't have to make do with existing materials and conform to the constraints they impose on us. To put it differently, in making actions, we bear a relation to the world different from the one we do in making physical artifacts. Given such a difference, there is no reason to think that the kinds of matter used in these two cases must be the same.

I am not convinced by Eynine's response. If the term "matter" has two radically different connotations, wouldn't it be more defensible to claim that there are two matter relations, one of them involving artifacts and the other involving actions? But if so, why should we literally take actions to be artifacts, as Eynine does? Furthermore, Eynine's response undermines a main advantage of amorphic hylomorphism. Amorphic hylomorphism aims to counter reductionism; it aims to tell us why there are tables, not just

particles arranged table-wise.<sup>27</sup> Here's Eynine's answer: "we have made that further object. We took the bricks and mortar and made a house out of them. We took the bronze and made a statue out of it" (2016, 18). But if the matter of an action is freely created by us, in what sense do we make the further action? We do not take certain bodily movements and make an action out of them; we simply create an action that involves the bodily movements. Wouldn't this make the action just "bodily movements arranged action-wise"? If so, the price Eynine has to pay to make his response work is to make amorphic hylomorphism appear unmotivated.

My second worry concerns the unifying function Eynine attributes to intentions. In cases where one clearly knows what one intends to do and how one intends to do it, one's intention seems capable of unifying the various actions one performs. There are nevertheless many cases in which such knowledge is lacking. Suppose Carl notices a new book in a bookstore. He skims through the book and finds it very entertaining. He wants to buy the book, but he has already spent too much money on books this month. So he begins to debate with himself, while picking up the book and putting it down several times. Eventually he makes up his mind and takes the book to the cashier. I don't see any problem with viewing the entire process as a single action, which may be described as the action of thinking about whether to buy the book or that of persuading oneself to buy the book. But is there any overarching intention in this case? The best candidate is the intention to think about whether to buy the book, but it is easy to describe a scenario in which Carl doesn't

---

<sup>27</sup> The term "things arranged tablewise" comes from Peter van Inwagen. See van Inwagen (1990, 110).

even have such an intention. Suppose that each time when Carl is about to put the book down, he feels ready to go home without the book. At these moments, his only intention is to go home. Whenever he actually puts the book down, however, it strikes him how hard it is to find a book one really likes. He then feels compelled to pick up the book once more and begins again to entertain the intention of buying it. Hence, while there is an interplay of the intentions of going home and buying the book, there is never the intention to think about buying the book. Despite the unity of the actions, no overarching intention can be found.

One could certainly deny that the actions feature any unity. Nevertheless, while it is plausible to hold that the process can be divided into smaller units, I find it hard to believe that these smaller units do not in any sense make up a larger action. Going back and forth is typical of any kind of deliberation. If one refuses to grant any unity to the actions in our example, one has to hold the same position with regard to most deliberative actions. Such a conception of actions seems unnecessarily restrictive. Moreover, it is hard to see how there can be any collective actions on this picture. Since there is a clear sense in which any collective action results from a plurality of individual intentions to bring about an action with other people, collective actions degenerate into “individual actions arranged collectively”, to use a phrase similar to van Inwagen’s. No collective actions ever feature the kind of unity individual actions feature. This worry is in fact made clear by one of Evidentiary’s own examples, in which he describes a collective attempt to fell a tree (2016, 236–37). A few people work in shifts to cut down a tree, but the shift schedule is flexible. When a person is tired, there are multiple people that can take over. This case is offered as an example of actions that do not

unfold according to a predetermined plan.<sup>28</sup> To further highlight my worry, we could stipulate that some of these people are training their skills while the others are just doing it for fun. There isn't a set time before which the people must fell the tree – it doesn't even matter if they stop before any tree is cut down. These people lack not only an overarching plan but also a consensus on the goals that must be completed. Given that the individual intentions do not overlap, nothing like a collective intention exists in this example. Even so, there is little doubt that the people are engaging in a collective action, that of felling a tree. It then becomes clear that many actions, individual or collective, are not unified by intentions. There isn't any intention that is entertained throughout the whole process. Again, the CP problem remains an unresolved issue.

### 3.4 Jones' Account

The third hylomorphic theory of events we shall consider is that of Tessa Jones. Her theory is formulated within the framework of Baker's theory of material constitution, so it is instructive to begin by considering how Baker characterizes the constitution relation. Just like Eynine, Baker (2007) takes the relata of the constitution relation to be numerically distinct entities. Therefore, if *The School of Athens* is constituted by a quantity of lime plaster, the fresco and the lime plaster are numerically distinct. The relata of the constitution relation must fall under different *primary kinds*, where "A primary kind is a kind in virtue of which a thing has its persistence conditions" (2007, 35). For example, while BLUE WHALE is a primary

---

<sup>28</sup> Eynine certainly doesn't intend this example to be a counterexample to his theory; he invokes it while discussing whether the matter of an action may change in the way the matter of a physical artifact does. He thus seems unaware of its possible implications.

kind, ENDANGERED ANIMAL is not. It is in virtue of being a blue whale, and *not* in virtue of being an endangered animal, that a blue whale counts as continuing to exist after, say, being lightly wounded in the ocean. Therefore, though blue whales would continue to exist even if they cease to be endangered thanks to new recovery plans, they could never exist without being blue whales. Just like BLUE WHALE, both LIME PLASTER and FRESCO are primary kinds. For this reason, a fresco may be constituted by a quantity of lime plaster.

For an entity to constitute another, the two must be suitably related. Say that an entity of primary kind  $F$  is in  $G$ -favorable circumstances if it meets a set of conditions  $C$  such that (a) the satisfaction of  $C$  by any entity of primary kind  $F$  guarantees the existence of an entity of primary kind  $G$ , but (b) the satisfaction of  $C$  by other entities may not (2007, 160). For example, one could create fresco-favorable circumstances for a quantity of lime plaster if one smears it on a wall and applies colors on it while it is wet, etc. But these circumstances may fail to bring about a fresco when entities of other primary kinds, such as a quantity of molten lava, are put in them. Baker defines constitution in terms of favorable circumstances. Take  $x$  and  $y$ , which fall under primary kinds  $F$  and  $G$ , respectively. Then  $x$  constitutes  $y$  if  $y$  is the entity that comes into existence when  $x$  is put in  $G$ -favorable circumstances (2007, 161). Note that this is the case only if certain constraints are not violated. First,  $y$  must occupy the same spatial location as  $x$ . A quantity of steel in Tokyo could never constitute the London Eye. Second, the basic kinds of stuff that characterize  $x$  and  $y$  must be the same. As Baker notes, the constraint is posited “to preclude the constitution of any immaterial thing by a material thing” (2007, 165). Third, the existence of an entity of primary kind  $F$  must not itself guarantee the existence of an entity of primary kind  $G$ . While the existence of a British



citizen is sufficient for the existence of an English speaker, British citizens can hardly be said to “constitute” English speakers. They *are* English speakers.

As an example, consider again *The School of Athens* and the quantity of lime plaster co-located with it. They do not violate any of the constraints. Hence, once the lime plaster is put in fresco-favorable circumstances, an entity of the primary kind FRESCO is created. In our example, this entity just is *The School of Athens*. It follows that *The School of Athens* is constituted by the lime plaster. According to Baker, though the two entities are numerically distinct, they are not *separate*: they share many properties on the basis of the constitution relation (2007, 169–70). The lime plaster would not have the property of being representative of the Renaissance style of art if it did not constitute *The School of Athens*. However, it doesn’t follow that the lime plaster fails to exemplify this property: no one would deny that a work of art representative of the Renaissance style would be destroyed if the lime plaster is destroyed. In Baker’s analysis, we should say that the lime plaster has the property, though it does so *derivatively*. By contrast, *The School of Athens* has the property *non-derivatively*; to be representative of the Renaissance style, *The School of Athens* does not have to constitute, or be constituted by, any other entity (See Baker 2007, 166–69). If Baker’s theory is correct, numerically distinct entities can acquire properties from each other. It would then be a mistake to view the fresco and the lime plaster as completely unrelated entities.

Jones (2013) argues that the conceptual resources provided by Baker also enable us to make sense of event constitution. One of her favorite examples is the ratification of the U.S. Constitution. Consider an event that is a signing of the Constitution. If such an event occurs in ratification-favorable circumstances, an event that is a ratifying of the Constitution occurs. The signing event can then be said to constitute the ratifying event (2013, 83). Jones

doesn't say much about the ratification-favorable circumstances, but they are easily identifiable: the authority of the signers must be acknowledged by the public and the need for a constitution must have been agreed upon, etc. One complication in applying Baker's theory to events is that an event may constitute another even if the two events unfold over slightly different time intervals. When the U.S. Constitution was signed in the Constitutional Convention, the ratification process was not yet over; it also had to be accepted by the legislatures of the states. It may be tempting to think that the events occurred successively and seek to understand their relation as something other than constitution (2013, 79–80). In Jones' view, the main alternatives are supervenience and causation. She thus examines and argues against these alternatives.

Jones attributes the supervenience-based view to Kim and offers counterexamples to it. Her example is the possible event in which Brutus failed to stab an immortal Caesar to death (2013, 74–76); I prefer a more mundane example. Suppose Anton's hands undergo a pattern of movement that counts as an expression of "the bus left" in American Sign Language (ASL). What is the relation between Anton's hand movement and the ASL utterance, if not constitution? One possible answer is that the utterance supervenes on the hand movement: if the properties of the utterance were different, the properties of the hand movement would be different. But this is false. In the actual world, Anton's interlocutor understands ASL, but there may be possible worlds in which she does not. It follows that the utterance could lose the property of being a successful linguistic exchange while co-occurring with the same hand movement. If so, the utterance does not supervene on the movement. Would it make more sense to say that the movement causes the utterance? Jones thinks that such an analysis is incompatible with Lewis' counterfactual understanding of

causation (2013, 80–81). We briefly discussed Lewis’ conception of causation in the previous chapter.<sup>29</sup> In the present example, for the utterance to causally depend on the movement, the utterance must occur if the movement were to occur. That would not be the case in many scenarios: Anton could coincidentally move his hands the same way in a world where ASL doesn’t exist. Therefore, if causation is understood counterfactually, it is implausible to treat the present case as one of causation.

I agree with Jones that hylomorphism offers the best analysis of the kind of phenomena she is interested in, but I am not satisfied with Jones’ version of hylomorphism. It has two main problems. First, Jones tries to dispense with Baker’s notion of primary kinds, and this makes her account vulnerable to counterexamples. Second, just like the others, Jones is unable to successfully respond to the CP problem. Let’s begin with the first. Jones attributes a single primary kind to all events: “The primary-kind property of an event is to have happened” (2013, 78). A congressional hearing and a tennis game therefore fall under the same primary kind: they are both entities that have occurred. It is a short step from the idea that all events are of the same primary kind to the conclusion that primary kinds contribute nothing to the constitution of events. This is indeed the conclusion Jones draws (2013, 78 fn. 1). As a result, she dispenses with primary kinds in her official account and elects to invoke only kinds in general.

Jones’ decision has unfortunate consequences. Take her example of the signing of the U.S. Constitution. This event certainly falls under the event-kind *signing of the U.S. Constitution*, but this is not the only kind the event falls under. It also falls under such kinds as

---

<sup>29</sup> See section 2.4.3.

*events occurring in September of 1787, events witnessed by George Mason, and indirect causes of the establishment of the U.S. Supreme Court in 1789.* However, if all the kinds an event falls under may play a role in constituting another event, we are forced to say that sentences describing the constitution relation feature intensionality. If the signing of the Constitution bears the matter relation to the ratification of the Constitution, the reason is presumably *not* that the signing is an event witnessed by George Mason. Hence, when viewed one way, the signing event constitutes the ratification event, but the signing event fails to do so when viewed another way. It is unclear to me that there are linguistic data supporting such a claim about intensionality. And even if there are, I am not sure how to make sense of the metaphysics. Isn't it the case that an event either constitutes another or it doesn't? From what Jones actually says, it is unclear that she has a response to this worry. In specifying what must be the case for event  $x$  to bear the constitution relation to event  $y$ , Jones employs sentences that quantify over "any event of  $x$ 's kind" and events "of the kind of  $y$ " (2013, 83). Unfortunately, there are no kinds that can be regarded as  $x$ 's kind *simpliciter* or *the* kind of  $y$ . The only candidates for such kinds are the primary kinds of  $x$  and  $y$ , but Jones is unwilling to incorporate them into her theory. I am not sure what justifies her position.

As a rejoinder, the friend of Jones' view could say that  $x$  and  $y$  stand in the constitution relation insofar as one of the kinds  $x$  falls under is appropriately related to one of the kinds  $y$  falls under. This won't do, either. Jones' view that all events fall under the same primary kind entails that the only thing we can say about the necessary properties of an event is that it is necessarily something that has occurred. So long as an event has indeed occurred in the actual world, it doesn't matter whether it falls under completely different event-kinds in other possible worlds. Hence, even if the actual signing of the U. S.

Constitution had not been a signing event, it could have been the very same event. It then transpires that for the suggested rejoinder to work, only the kinds that an event falls under *in the actual world* should be considered. For given any kind  $K$  under which an event falls, there will be possible worlds in which the event doesn't fall under  $K$ . To make the decision on whether  $K$  is one of the kinds under which the event falls – a decision required by the suggested rejoinder – the most natural choice is to focus on the actual world. But now Jones' account is faced with something similar to one of the problems plaguing Crowther's account. Take Fine's example of an arm-raising action. If we only focus on the event-kinds this event falls under in the actual world, then it certainly falls under the kind that defines the corresponding bodily movement. What reason is there to say that there are two events, an action and a bodily movement? None, I would think. If so, hylomorphism about events appears unmotivated; we should simply go for the more intuitive theories of events.

Let's now turn to my second objection: Jones' view doesn't offer a solution to the CP problem, either. Like Crowther and Eynine, Jones acknowledges the possibility of constitution by a plurality. Her example is a wedding constituted by the married couple's vows, promises and other activities. Jones argues that the collection of events doesn't feature a primary kind – in fact, what she says makes it sound as if the collection didn't fall under any kind: “If it didn't come to constitute the wedding, the collection would be like any other collection of events and just be a random sum, with the only thing essential to it being that it has the same members” (2013, 78 fn. 1). The only event-kind the collection falls under is the kind RANDOM SUM, if that counts as a kind at all. But such a view makes it inexplicable why the various events in the collection constitute any single event. To show that they do, we have to identify the wedding-favorable circumstances. Are there circumstances under which

the existence of a random sum of events is sufficient for the occurrence of a wedding? I cannot think of any. The following conditional certainly holds: if a random sum of events exists, then some events occur. But the application of Jones' theory depends on the truth of a much stronger conditional that has the following form: if a random sum of events exists, then there is a kind *K* such that some events of kind *K* occur. I don't believe that any conditional of such a form is true, let alone the one in which we substitute the kind WEDDING for *K*.

But maybe Jones' theory still has the resources to account for constitution by a plurality despite the problematic stance she officially takes on this issue. Maybe we should admit primary kinds into our theory again and say that even the collection of the events supposedly constituting the wedding falls under a primary kind. To see how this might help, consider again the case of the no-confidence vote discussed in section 2.4.3. We have used the name **sum** to call the mereological fusion of **speech** and **vote**, and we have used the name SUM to call the individuating property of **sum**. One could argue that SUM corresponds one-to-one to a kind, i.e. the kind *a vote followed by a speech* (henceforth abbreviated as V&S). Let's take V&S to be the primary kind of **sum**. The following conditional seems true: if an event of the kind V&S is in circumstances favorable to a successful no-confidence motion, then there is an event of the kind under which all successful no-confidence motions fall. Doesn't this show that **sum** constitutes **no-confidence**? If so, why don't we treat the wedding similarly and take something like the kind *vows followed by promises followed by blessings* to be the primary kind of the wedding-constituting collection of events?

We should not accept this response if we take into consideration Jones' original insight. She is right in saying that a collection of events doesn't feature any primary kind that

is relevant to the phenomenon of constitution: “The primary-kind property that collection has is assigned because it constitutes a wedding” (2013, 78 fn. 1). It may be true that **sum** falls under the kind V&S. Even so, one shouldn’t make the further claim that V&S is a kind responsible for the persistence conditions of all the entities that fall under it. If the primary minister’s speech counts as continuing to unfold after a cough by the prime minister, it is because the speech falls under the kind defined by SPEECH, not because the speech falls under the kind V&S. To claim that there are events that derive their persistence conditions from V&S is to admit entities like the fusion of my brain and the London Eye into the realm of events. As stated in Section 2.2, I believe that we shouldn’t accommodate such entities in our ontology. If my reasoning is correct, then at the end of the day, Jones’ account doesn’t fare better in solving the CP problem than the other hylomorphic theories on offer.

### **3.5 An Alternative Hylomorphic Account**

#### *3.5.1 Fine and Johnston on Event Constitution*

In this section, I shall defend an alternative hylomorphic account of events. Since some of the basic ideas in my proposal can be found in the work done by Kit Fine and Mark Johnston, it is instructive to briefly examine their work. We have already seen that Fine motivates a hylomorphic theory of events by arguing that an action differs modally from the bodily movement co-occurring with the action. The action is performed with a certain intention, and the property of being so performed supplies the action with a form. It may seem that Fine’s theory is only applicable to intentional actions, but he intends it to apply to events in general. To show how this is done, Fine discusses Sir Lancelot’s loss of virginity

(1982, 103–4). Suppose it happened at  $t$ . Had Sir Lancelot had sex before  $t$ , the same physical activity could still have happened at  $t$ , but it would have not been a losing of virginity. It follows that the loss of virginity and the physical activity were modally distinguished, even though they coincided spatiotemporally. There was a form present in the latter but absent from the former: the property of preceding any other sexual activity. These considerations justify speaking of the physical activity as constituting the loss of virginity. Since intentions do not figure in these considerations, they support a hylomorphic view of events in general. Fine even goes so far as to claim that a single class of events, *occurrences*, constitute all the other events: “Occurrences are the matter of events; and just as all material things are ultimately formed from matter... so are all events ultimately formed from occurrences” (1982, 104).

Johnston has a similar view about events. On his view, the matter of an entity consists of its parts, while the form is the principle of unity that holds the parts of the entity together (2006, 676). Since parts are construed broadly, the successive phases of an event count as its parts. These parts stand in various relations to each other, and some of the relations may function as principles of unity. In Johnston’s example of a dinner party, the principle is “occupying a position in a sequence of events each effected either by the hosts or by the guests for the intended outcome of their mutual feeding, entertainment, and relief” (2006, 657; emphasis omitted). Since relations are usually taken to be polyadic properties, Johnston’s view resembles Fine’s in that forms are regarded as properties. However, Johnston’s view features an element not present in Fine’s view: the constitution relation holds not only between events but also between physical objects and events. For example, Hannah’s serve in a tennis game might be thought of as an event constituted by Hannah, the



tennis ball, the relation of hitting and the spatiotemporal region in which the game takes place. These entities constitute a serving event if they are governed by the following principle of unity: Hannah bears the relation of hitting to the ball at the region. In Johnston's terminology, this principle of unity is *item-generating*: it takes certain extant items as input and generates a new item, which is an event. Principles of unity that generate events are called *event-builders* (2006, 676–77).

The main issue with Fine's and Johnston's views is that they do not delve into details. They give us some general ideas that need fleshing out. Their central insight is that principles of unity should be understood as properties: a hylomorphic entity is more than its matter because it exemplifies properties that unify its matter. They are nevertheless silent on how exactly properties unify matter. Fine speaks of a historical property that unifies what Sir Lancelot does in a sexual encounter and generates his loss of virginity. But what's so special about this historical property that it can play the unifying role in the particular context? What exactly is the relation between this historical property and the sexual activity? Saying that it is the matter relation or the constitution relation amounts to nothing more than naming the relation. Similarly, in trying to show what unifies a dinner party, Johnston cites a polyadic property whose components include temporal, causal, agential and social properties. This example barely gives us any hints on how we should go about looking for a general solution to the CP problem. In short, Fine and Johnston haven't really identified the conditions on which an event or a set of events constitutes another. As things stand now, their views are not well-developed alternatives to the theories proposed by Crowther, Evnine and Jones, which appeal to temporal boundaries, intentions and favorable circumstances, respectively.

Therefore, in the next section, I shall develop a version of hylomorphism broadly based on Fine's and Johnston's views and work out some of the needed details.

### 3.5.2 *Dependence-Based Hylomorphism*

As I have tried to show, the CP problem is one of the main challenges faced by theories of events. Since my proposal is motivated by the need to respond to this problem, let's address it head-on and return to the case of the no-confidence motion once again. What is the relation **no-confidence** bears to **speech** and **vote**? Obviously, the answer I prefer is that **speech** and **vote** jointly constitute **no-confidence**; they bear the matter relation to **no-confidence**. But what are the conditions met by **speech** and **vote** such that they jointly constitute **no-confidence**? My answer, as I briefly suggested in the introductory section, is that they jointly supply what is needed for **no-confidence** to go on. To put it differently, they succeed one another in affording the means by which **no-confidence** continues to unfold. The event that is **no-confidence** is initiated precisely because **speech** occurs; it is in virtue of the prime minister's speech that a passing of a no-confidence motion comes into being. In a similar fashion, when **speech** is over, **no-confidence** continues without interruption because **vote** occurs; the passage of the no-confidence motion is able to run its course in virtue of the collective action of the parliament members to vote in response to the prime minister's speech. In this picture, the relation **speech** and **vote** collectively bear to **no-confidence** is analogous to the relation the matter of the London Eye bears to the London Eye itself. Just as we can destroy the London Eye by destroying the quantity of metal it is made of, we can prevent **no-confidence** from occurring by preventing **speech** and **vote**

from occurring. We are thus justified in saying that **speech** and **vote** bear the constitution relation to **no-confidence**.

What I have said is intended to be an intuitive characterization of my answer to the question: what are the conditions met by **speech** and **vote** such that they jointly constitute **no-confidence**? Now we need to make the answer more precise. It would seem that what I have said can be summarized in a single sentence: events  $e$  and  $f$  constitute another event  $g$  if and only if the successive occurrence of  $e$  and  $f$  is sufficient for the occurrence of  $g$ . This obviously has to be qualified. It doesn't take much reflection for one to realize that the successive occurrence of **speech** and **vote** is not really sufficient for the occurrence of **no-confidence**: if the parliamentary democracy under consideration had different laws, there could be **speech** and **vote** without **no-confidence**. In another possible world, the law could say that a no-confidence motion cannot be passed without a referendum; in that world, **speech** and **vote** only mark the beginning of a successful no-confidence motion. To be sufficient for **no-confidence**, **speech** and **vote** must take place while certain laws are in place. Should we then say that **no-confidence** is constituted by **speech**, **vote** and the relevant laws? No, because the same problem would arise again. The relevant laws would lose their binding power if law enforcement did not exist, so law enforcement too must figure in the sufficient conditions of **no-confidence**. Needless to say, it would be untenable to add law enforcement to the growing list of the entities by which **no-confidence** is constituted. While we might be able to stop extending the list after finitely many items are added to it, taking constitution to be a relation with such a huge number of relata renders any theory of constitution unintelligible.

The way out of this impasse, I submit, is to invoke a distinction drawn by Husserl. As Husserl points out, the dependence relation between two entities can be more or less immediate ([1913b] 1970, 2:28).<sup>30</sup> Suppose Bernard was born with a rare disease whose lethal symptoms could only be suppressed by a medication patented by a pharmaceutical company. Since the company could not exist without its shareholders, Bernard could not exist if none of the shareholders did. While it is true that Bernard depends on the shareholders, he does so only because he depends on the medication. We could then say that the dependence relation between Bernard and the medication is *immediate*, but the dependence relation between Bernard and the shareholders is *mediate*. The same analysis can be applied to events. While **no-confidence** immediately depends on **speech** and **vote**, its dependence on the relevant laws, law enforcement and everything else is mediate. We can then refine our criterion: events *e* and *f* constitute another event *g* if and only if *g* immediately depends on nothing but the successive occurrence of *e* and *f*.

Why is the dependence of **no-confidence** on **speech** and **vote** immediate? The reason is that they are the means by which **no-confidence** occurs. Without such means, nothing whatsoever would take place. In other words, in a world where the prime minister gives no speech and the parliament members fail to cast their votes, **no-confidence** is not only deprived of some of its properties; it is deprived of its very eventhood, thus its very existence. In contrast, in a world where the relevant laws aren't in place, **no-confidence** could still occur, though it would lose its legal properties and some other properties. Even if

---

<sup>30</sup> The context of Husserl's analysis of dependence is his theory of *foundation*, where foundation should be regarded as a specific kind of dependence. For discussions of Husserl's theory of foundation, see Simons (1982), Fine (1995) and Correia (2004).

one is reluctant to say that the event in that world is numerically identical to **no-confidence**, it can hardly be denied that some event of a similar kind would still exist. That event would still share many social and political properties with **no-confidence** and occupies exactly the same spatiotemporal location as **no-confidence**. This is not the case for a world where no speech and vote take place – in such a world, no events capable of bearing any property would exist at the spatiotemporal location of **no-confidence**. It is thus reasonable to say that **no-confidence** depends on the relevant laws only because it depends on **speech** and **vote**. Since the same can be said of any other entity upon which **no-confidence** depends, the dependence of **no-confidence** upon **speech** and **vote** is immediate.

But we have to be careful about the kind of dependence we are talking about here. It is not true that **no-confidence** immediately depends on the very event tokens of **speech** and **vote**. It is possible for **no-confidence** to occur even if **speech** and **vote** did not. Suppose the prime minister wrote the speech herself but got sick before delivering it. As a result, a member of the cabinet steps in as her proxy and reads the speech word-for-word in front of the parliament members. This speech is presumably numerically distinct from **speech**, but I see no reason to hold that the event as a whole is numerically distinct from **no-confidence**. The speech so delivered is almost the same, so the legal and political properties of the no-confidence motion all remain unchanged.<sup>31</sup> Similarly, there are different ways the members of the parliament could vote on the no-confidence motion. They could vote by raising their hand instead of pushing a button, and that would probably result in an

---

<sup>31</sup> Many would probably remain unconvinced if I could not offer any general observations on how different an event could have been without amounting to a different event. I will say more about that in Section 3.6.2.

event different from **vote**. But again, I don't think such a difference compels us to say that the event as a whole differs from **no-confidence**. It then becomes clear that **no-confidence** immediately depends on event tokens of the kinds under which **speech** and **vote** fall, respectively, but it doesn't immediately depend on **speech** and **vote**. Our criterion thus needs further refinement.

As a preliminary to formulating our criterion precisely, let's follow Peter Simons in distinguishing between two different dependence relations (1994, 559–60, 567):

(Individual Dependence)

Let  $x$  and  $y$  be arbitrary entities. Then  $x$  individually depends on  $y$  if and only if  $y$  exists in every possible world that contains  $x$ .

(Specific Dependence)

Let  $J$  and  $K$  be kinds. Then entities of kind  $J$  specifically depend on entities of kind  $K$  if and only if there is some  $y$  of kind  $K$  in every possible world that contains some  $x$  of kind  $J$ .<sup>32</sup>

For instance, there would be no democracies without electoral systems. It follows that any possible world containing a democratic country also contains an electoral system. By the above definition, entities of the kind DEMOCRACY specifically depend on entities of the kind ELECTORAL SYSTEM. Since Estonia is a democracy, we could say that Estonia specifically depends on an electoral system. Even so, Estonia doesn't individually depend on its current

---

<sup>32</sup> The term "specific dependence" means, in Simons' own words, "dependence at the species level" (1994, 560). Simons' account of dependence is based on his interpretation of Husserl's theory of foundation. For additional conceptual and formal details of Simons' interpretation of Husserl, see his (1982).

electoral system. Estonia would continue to exist even if Estonians were to change the current system and bypass the parliament in presidential elections, for example.

To resume our discussion of event constitution, we should take the dependence of **no-confidence** on **speech** and **vote** to be specific rather than individual. The reason **no-confidence** depends on the two events is that it falls under the kind under which all and only successful no-confidence motions fall.<sup>33</sup> In any possible world where a successful no-confidence motion occurs in the given country, a speech by the prime minister and a vote by the parliament members also occur. However, as we have seen, the speech and the vote need not be numerically identical to **speech** and **vote**. It follows that successful no-confidence motions specifically depend on the kinds under which **speech** and **vote** fall. To describe our case in Simons' terminology, we could say that **speech** and **vote** satisfy the *dependence needs* of **no-confidence** (P. Simons 1994, 562). Unless otherwise noted, I shall use the term to denote *needs for immediate specific dependence*, so to say that the dependence needs of **no-confidence** are satisfied by **speech** and **vote** is to say the following: on the one hand, the dependence of **no-confidence** on **speech** and **vote** is specific and immediate, and on the other hand, **no-confidence** only bears such a dependence relation to **speech** and **vote**. Now that we have defined dependence needs, we can revise our criterion of event constitution as follows. Suppose events  $e$  and  $f$  fall under kinds  $E$  and  $F$ . Then  $e$  and  $f$  constitute another event  $g$  if and only if there is a kind  $G$  such that  $g$  falls under  $G$ , on the one hand, and the

---

<sup>33</sup> More precisely, the tokens of the kind include all and only the no-confidence motions that count as successful ones according to the actual legal system of the country under consideration.

dependence needs of any token of *G* are satisfied by both a token of *E* and a token of *F*, on the other.

Before we can present an official formulation of our theory, additional comments are in order. First, my proposal explicitly invokes event-kinds to define event constitution. As things stand now, my proposal is confronted with the same difficulty confronting Jones' theory: most, if not all, events fall under multiple kinds. Since dependence needs have their origins in kinds, the dependence needs an event has vary with our choices of the kinds to focus on and describe. What is even worse is that *whether* the dependence needs of an event are satisfied also becomes a description-relative matter. To avoid this problem, I shall make more substantial use of Baker's idea of primary kinds. *Pace* Jones, events do not share a single primary kind. In our example, **no-confidence**, **speech** and **vote** fall under distinct primary kinds. The dependence needs of an event originate in the primary kind the event falls under. Second, as we have seen at various points, constitution is supposed to a relation between spatiotemporally coinciding entities. Hence, we should require **speech** and **vote** to jointly occupy the same spatiotemporal location as **no-confidence**. Note that we need not make the assumption that **speech** and **vote** have a fusion; all we need to assume is that the spatiotemporal regions they occupy have a fusion. While the former assumption may be problematic, the latter seems reasonable. Third, it may be a little pedantic to point out that an event may fail to exist even if its needs for immediate specific dependence are satisfied, but mentioning this point may help to guard against misinterpretations. For an event to exist, everything it depends on must exist, not just those it depends on immediately and specifically. While those entities on which an event does not immediately and specifically depend do not play a role in constituting the event, they do play a role in the existence of the



event. After all, if something on which an event mediately depends fails to exist, then presumably the entities on which the event immediately depends fail to exist.

With these points made, I can officially state my version of hylomorphism, which I call *dependence-based hylomorphism* (henceforth DBH). My point of departure is Fine's view that there is a class of *occurrences* without which no events could be constituted. Given this view, it makes sense to divide events into simple and complex ones, where simple events just are Fine's occurrences.<sup>34</sup> Fine doesn't say much about the nature of occurrences, but I think it is plausible to identify them with tropes of a certain kind. By combining hylomorphism with a trope theory of events like Bennett's,<sup>35</sup> we get the best of both worlds: we are able to say something about the nature of simple events, which Fine doesn't do, and we can explain how various simple events constitute a complex event, which Bennett fails to do. For this reason, I shall assume that the class of simple events consists of tropes, though I don't have a criterion we can use to identify these tropes. Complex events are related to simple events by the constitution relation. A complex event is constituted by a number of simple events just in case that the dependence needs of the former are satisfied by the latter. The simple events are the matter of the complex event, but there is more to the complex event than its matter. It also has a form, which is derived from its primary kind. In our example, the form of **no-confidence** is the following property: being the kind under which all and only successful no-confidence motions fall.<sup>36</sup> By identifying **no-confidence** with the composite

---

<sup>34</sup> My distinction of simple and complex events is stipulative. I do not mean to suggest that this is the only way to draw the line between simple and complex events.

<sup>35</sup> See section 2.4.3.

<sup>36</sup> Baker uses the term *primary-kind property* to denote an entity's property of falling under its primary kind; see Baker (2007, 33–34).

that has this property as its form and **speech** and **vote** as its matter, we are also able to define its existence and identity conditions: **no-confidence** exists if its dependence needs are satisfied, and it is numerically identical to any event that has the same form and matter.

Putting everything together, we can formulate the five components of dependence-based hylomorphism, DBH I-V, as follows:<sup>37</sup>

(DBH I: The Satisfaction of Dependence Needs)

Let  $x, y$  and  $z$  be events whose primary kinds are  $I, J$  and  $K$ . We say that  $x$  and  $y$  satisfy the dependence needs of  $z$  if and only if:

- (1) for any event  $k$  of primary kind  $K$ , there are events  $i$  and  $j$  of primary kinds  $I$  and  $J$  such that  $k$  specifically and immediately depends on  $i$  and  $j$ ;
- (2) the fusion of the spatiotemporal regions occupied by  $x$  and  $y$  is numerically identical to the spatiotemporal region occupied by  $z$ .

(DBH II: Direct and Indirect Constitution)

Let  $x, y$  and  $z$  be events. Then  $x$  and  $y$  constitute  $z$  if and only if:

- (1) EITHER  $x$  and  $y$  directly constitute  $z$ , where this is the case if and only if:
  - (a)  $x$  and  $y$  satisfy the dependence needs of  $z$
  - (b) everything else  $z$  depends on also exists;
- (2) OR  $x$  and  $y$  indirectly constitute  $z$ , where this is the case if and only if there is a series of events such that:

---

<sup>37</sup> To make the definitions readable, I focus on the cases in which two events jointly constitute another. But it should be clear that the definitions can be used to analyze the constitution of an event by an indefinite number of others.

- (a)  $x$  and  $y$  directly constitute the first event in the series;
- (b) every event in the series directly constitutes the next;
- (c) the last event in the series directly constitutes  $z$ .

(DBH III: Form and Matter)

If events  $x$  and  $y$  constitute event  $z$  and  $z$  is of primary kind  $K$ , then we say that:

- (1)  $x$  and  $y$  are the matter of  $z$
- (2) the property of falling under  $K$  is the form of  $z$ .

(DBH IV: Event Existence)

Any event  $x$  is either simple or complex.

- (1) If  $x$  is simple, then  $x$  exists only if  $x$  is a trope.
- (2) If  $x$  is complex, then  $x$  exists if and only if there are some simple events that constitute  $x$ .

(DBH V: Event Identity)

Any event  $x$  is either simple or complex.

- (1) If  $x$  is simple, then for any simple event  $y$ ,  $x$  is numerically identical to  $y$  if and only if  $x$  and  $y$  are the same trope.
- (2) If  $x$  is complex, then for any complex event  $y$ ,  $x$  is numerically identical to  $y$  if and only if:
  - (a) the primary kinds of  $x$  and  $y$  are the same;
  - (b)  $x$  and  $y$  are constituted by the same events.

These summarize everything I have said in this section. To see how it works, let  $x$  and  $y$  be **speech** and **vote** and  $z$  **no-confidence**. By DBH I-II, **speech** and **vote** directly constitute **no-confidence**. By DBH III, the matter of **no-confidence** consists of **speech** and **vote**, and its form is the property of falling under the kind under which all and only successful motions of no-confidence fall. By DBH IV, **no-confidence** exists. By DBH V, **no-confidence** is numerically identical to any successful motion of no-confidence that is constituted by **speech** and **vote**. DBH I-V therefore capture what I meant when I described **speech** and **vote** as affording what was needed for **no-confidence** to unfold.

Dependence-based hylomorphism has several virtues. First, it is specifically designed as a response to the CP problem. Though Crowther, Eynine and Jones all acknowledge the possibility of constitution by a plurality, I have argued that none of their theories explains how exactly constitution by a plurality is possible. DBH solves this problem. In contrast with Crowther's temporal sortals and Eynine's intentions, what unify events in DBH are dependence relations based on primary kinds. The latter are much more flexible than the former. Furthermore, DBH is able to circumvent the thorny issue of defining favorable circumstances for random collections of events, which plagues Jones' theory.

Second, DBH fleshes out the details missing in Fine's and Johnston's accounts. We are told by Fine that properties function as forms that turn occurrences into events. Similarly, Johnston tells us that relational properties bind individuals and spatiotemporal regions together into events. But what kind of properties are we talking about, exactly? Why are such properties capable of functioning as forms? DBH suggests an answer to this question: the properties in question are those defined by primary kinds. They are capable of functioning as event-forms because the very persistence conditions of an event derive from

the event's primary kind. The primary kind of an event thus determines what must happen for the event to run its course, which means that the primary kind of an event determines its matter.

Third, DBH doesn't rely on any analogies between events and physical artifacts. We don't have to say that we work on bodily movements just as we work on physical matter; nor do we have to say that events are made of non-events just as physical artifacts are made of stuff. Since dependence-based hyломorphism does without these analogies, its strength and weakness can be evaluated in a more straightforward manner than Evinne's and Crowther's theories. This also means that DBH presupposes a simpler ontology. We don't have to posit such entities as non-events that in some sense correspond to quantities of physical stuff; tropes and properties are all we need. If one is willing to reduce properties to classes of tropes, then DBH in effect supplies supporting evidence of how much a trope ontology can do. Given these considerations, we have good reasons to accept DBH. Before the chapter is concluded, however, I would like to address two potential objections to any constitution-based view about events.

## **3.6 Objections and Replies**

### *3.6.1 Realization, Not Constitution*

A concern one might have with any constitution-based view about events is that it seems forced. Consider again Fine's example. Isn't it more natural to describe the relation

between the bodily movement and the action as that of *realization*?<sup>38</sup> Instead of saying that the bodily movement constitutes the action, we should say that the bodily movement realizes the intention to act in a certain way. Similarly, instead of saying that **speech** and **vote** constitute **no-confidence**, we should say that the successive occurrence of the former realizes the latter. Whether this alternative is preferable obviously depends on how the realization relation is construed. If one thinks that realization is a relation an entity bears to its matter, then the dispute may be a terminological one. We could speak of the London Eye as realized by the steel used to construct it; I have nothing against this manner of speaking.

But the dispute may be more than terminological. One could say that there are actually no actions in the scenario described by Fine. What can be found there are the bodily movement and the agent's intention. Since the bodily movement realizes the intention, we describe the movement as an action; such a description nevertheless picks out nothing but the bodily movement. Or take **speech** and **vote**. Since **speech** and **vote** jointly realize a political mechanism, we conceptualize what happens as the passage of a no-confidence motion. It is true that **speech** and **vote** exist; so does the political mechanism as an abstract entity. But everything else is in our head. On such a view, the relata of the realization relation are different from those of the constitution relation. The bodily movement either bears the constitution relation to an action or bears the realization relation to an intention. Since an action is not an intention and there are no reasons to posit unnecessary entities, the bodily movement does not bear the constitution relation to any action.

---

<sup>38</sup> I thank Qiong Wu for discussing this issue with me.

There are at least three ways to respond to this view. First, one could argue that **no-confidence** and the action in Fine's example are not unnecessary entities. As we have emphasized, they have different modal properties than the events that are their matter. The modal phenomena would become mysterious if **no-confidence** and the action were simply nonexistent. Second, it makes little sense to say that all constituted events are just in our head. We could ask: what would happen if **speech** and **vote** were to take place without the political mechanism in place? The natural answer is: in that case, there wouldn't be any successful motion of no-confidence. But on the view under consideration, that is the wrong answer. Instead, the answer should be: in that case, we have to understand the speech and the vote in a different way. The answer isn't wrong, but it sounds like a non-answer. We could follow-up with the question, "what makes it the case that we have to understand the speech and the vote in a different way?" At this point, any compelling answer would have to invoke a constituted event rather than something in our head. Third, simply giving a name to the relation between the intention and the bodily movement is far from giving a theory of the relation. Before such a theory is fleshed out, it is premature to assume that it is capable of explaining everything that a constitution-based view about events is capable of explaining.

It is true that there are richer conceptions of the realization relation on offer. As Karen Bennett points out, realization can be thought of as a relation between a causal role and an occupant of the role (2017, 10). Take the eye as an example. One could say that the eye is whatever causally mediates between light signals and neural activities in the optic nerves – any organ that plays this role realizes the eye. Hence, though an octopus eye is structurally different from a human eye, the former realizes the same role in an octopus as the latter does in a human being. The more we flesh out the idea of realization, however, the

less it seems the same as constitution. To say that an action is realized by a bodily movement, we have to identify the distinctive causal role played by the action. But in many cases of constitution, there doesn't seem to be any such role. Suppose an agent has the intention to swim in the pool in front of her, so she begins to jump. But she is suffering from a hallucination, so there is actually no pool in front of her. What events are supposed to be causally mediated by her jump? There are too many possibilities here: she could be tripped by something on the floor, hurt herself or find the action embarrassing, to name a few. The sheer number of possibilities suggests that the action is not individuated by any single causal role. If so, nothing can stand in the realization relation to the action. The same argument can be made about a wide range of constituted events.

The moral we should draw from these considerations, I submit, is that the constitution relation cannot be straightforwardly identified with the realization relation. While such identification remains a theoretical possibility, it needs to be supported by additional arguments. The arguments we have discussed so far do not threaten the theoretical usefulness of the constitution relation.

### 3.6.2 *Events Have No Contingent Properties*

I have suggested that in the aforementioned case of the no-confidence motion, **no-confidence** can occur in a possible world where **speech** and **vote** are replaced by slightly different events. In some of the worlds where a proxy of the prime minister steps in to deliver the speech and the parliament members choose to vote by raising their hands instead of using more advanced technology, we can still identify the very same token of **no-confidence**. Some might be unhappy with my claim. There are many ways to answer the



question: how many properties could an event have lost without turning into a different event? Some think that any change to an event gives rise to a distinct event. This is Crowther's view: "events cannot change their properties; they possess all of their properties necessarily" (2011, 36). If we accept this view, there are no worlds in which **no-confidence** occurs but the very tokens of **speech** and **vote** do not. To defend DBH, I have to show that this view is problematic.

An argument supporting my position can be found in Evinine's work. Evinine notes that there are many ways in which a single action could be carried out; there is no reason to think that one would have failed to perform a given action if the action had not been performed exactly the way it was actually performed (2016, 244). Suppose I just pressed a key on my laptop. If I had pressed the key slightly further down or struck it with a little less force, I would still have carried out the same action. Cases like this suggest that an event could have some properties contingently. I agree with Evinine's diagnosis. However, some might argue that we could also accommodate the cases by appealing to action types. We could hold that different ways to press the key amount to different actions, even though these actions belong to the same type. If we take this alternative, then we can stick to the view that all the properties of an event are necessary. Since intuitions about the modal nature of events are usually fuzzy, many might prefer this alternative. To defend the view that events have some properties contingently, I shall provide another argument.

My argument is a simple *reductio ad absurdum*: if we assume that every property of which an event is an instance is necessary to the event, then the actual course of history is the only possible course of history. Since this makes much of our everyday discourse unintelligible, we should reject the assumption. It is not hard to see how this argument can

be fleshed out. Suppose events can only instantiate properties necessarily. For example, if  $e$  is an instance of the Olympics in 2016, then  $e$  is necessarily an instance of the Olympics in 2016. It follows that everything that happened necessarily happened, everything that is happening now necessarily happens now, and everything that will happen will necessarily happen. Therefore, the course of history of the actual world is shared by every possible world. Some might think that this is too quick. They might argue that my reasoning is fallacious: even if  $e$  is necessarily the Olympics that occurred in 2016, it doesn't follow that there is any time at which  $e$  necessarily exists. Take Victor Frankenstein. Though Victor Frankenstein is necessarily a human being who lived in the 18<sup>th</sup> century, there isn't any time at which he necessarily exists. He has never existed and will never exist. Despite this, anyone who is not a human being living in the 18<sup>th</sup> century fails to fit the depiction in the novel, thereby failing to be Victor Frankenstein. Victor Frankenstein is thus a counterexample to my reasoning. It follows that even if the actual events could not possibly have been otherwise, many of them could still have failed to occur.

My response to this objection is that it rests on confusion. Suppose Victor Frankenstein has never existed and will never exist. Then he has no properties whatsoever, necessary or contingent – there is nothing to which the name “Victor Frankenstein” bears the referential relation, let alone something that has properties.<sup>39</sup> To make sense of the objection, we could instead take Victor Frankenstein to be a fictional character. A fictional character exists as an abstract object, but no human beings are abstract objects. It follows

---

<sup>39</sup> I have no intention to rule out the possibility that the name is still meaningful or that many speakers intend to use it as a referential expression.

that Victor Frankenstein is not necessarily a human being. As a rejoinder, my critics could take the approach described by Fabrice Correia: “It is nowadays widely accepted that to say that an object is essentially so and so is to say that necessarily, the object is so and so, or alternatively, that necessarily, the object is so and so *if it exists*” (2007, 63; my italics). My critics could reinterpret the sentence “Victor Frankenstein is necessarily a human being who lived in the 18<sup>th</sup> century” as a conditional, which says that if Victor Frankenstein had existed, he would necessarily have been a human being in the 18<sup>th</sup> century. This, however, doesn’t show that Victor Frankenstein is necessarily an instance of humanity in the 18<sup>th</sup> century. The conditional is true even if there is a possible world in which there were no human beings in the 18<sup>th</sup> century, and Victor Frankenstein cannot be an instance of humanity in the 18<sup>th</sup> century in that world. The only necessary property we can attribute to Victor Frankenstein on the basis of the conditional is the property of being human whenever he exists. Hence, the objection at most shows that even if Victor Frankenstein is necessarily a human being whenever he exists, it doesn’t follow that there is a time at which Victor Frankenstein necessarily exists. This is a trivial claim. What the objection fails to show is that even if Victor Frankenstein is necessarily a human being *living in the 18<sup>th</sup> century*, it doesn’t follow that there is a time at which Victor Frankenstein necessarily exists. There is obviously such a time: the 18<sup>th</sup> century. As it turns out, the objection doesn’t threaten the soundness of my argument.

This brings us back to the issue of the modal nature of events. The assumption in my *reductio* argument is that every property of which an event is an instance is necessary to the event. This includes such properties as being the Olympics *in 2016* or being a congressional hearing *today*, which are evidently properties an event could instantiate. As my

response to the potential objection suggests, the assumption implies that all the events in the actual world necessarily occur; moreover, they necessarily occur at the times at which they actually do. The actual world is, as it were, the minimal possible world; every other possible world must contain every event in the actual world. The only way for the non-actual worlds to differ from the actual world is for them to include events that are numerically distinct to every actual event. Furthermore, these events have to be causally isolated from the actual events. If this were not the case, the actual events would acquire additional causes and effects, which could easily lead to the problem of overdetermination. The logical consequence of our assumption, therefore, is that there are only two kinds of possible worlds. In possible worlds of the first kind, the actual course of history is the only series of events. These worlds might somehow differ from the actual world with respect to individuals, facts and fictional entities, etc., but nothing that doesn't happen in the actual world happens in these worlds. In possible worlds of the second kind, the actual course of history is one of the many series of events that are completely causally isolated from each other. A world of this kind is a weird one; perhaps it is constituted by spatiotemporal regions that are connected by wormholes.

I take this to be an extremely undesirable result. If these are the only two kinds of possible worlds, then much of our modal discourse is nonsense. One might say “the theft could have been easily avoided” or “the pollution could have been worse”, but the premise of our *reductio* argument entails that these are all falsities. An equally troubling consequence is that we will have to regard a seemingly plausible kind of historical explanation as wrongheaded. As Daniel Nolan points out, certain questions in the discipline of history demand answers that are naturally given in terms of counterfactual claims: “When we want

to know whether one outcome was very likely, we would naturally consider what sorts of events would have prevented it *if* they had happened, and how likely they were” (2013, 330). A concrete example can be found in the terrific book by Kellee S. Tsai (2007). Many share the idea that a country that has seen enough economic growth will eventually become a democracy. But there are authoritarian countries that seem to defy such a regularity, so historians have to explain why democracy hasn’t emerged in those countries. They look for such explanations by asking: what historical developments would have turned those countries into democracies if they had taken place? One of the theories can be stated with a counterfactual: these countries would have become more democratic if their economic system had been different and the middle-class citizens’ pursuit of material well-being didn’t depend so much on government cooperation (Tsai 2007, 22–23). This is a very sensible theory. However, if there is no sense in which past events could have taken different turns, then all counterfactuals are false. And no true theory consists of a set of false claims. We are faced with two options here: to reject the kind of historical explanation just mentioned, or to deny the premise that any property an event has is one of the event’s necessary properties. The more reasonable move, I submit, is to choose the second option. We thus conclude that events must have some contingent properties.

Does this mean that it is possible for an event to be radically different from the way it actually is? I don’t think so. To show that this other extreme is equally unattractive, it is useful to revisit Jones’ position. As we have seen, Jones maintains that an event could have been quite different. She claims that Brutus’ stabbing of Caesar, which killed Caesar, could have been a failed attempt to kill an immortal being (2013, 74). I disagree. I take it to be a tenable assumption that humanity was one of Caesar’s essential properties. Since humanity is

in part defined by mortality and Caesar was human, he couldn't have been identical to an immortal being – whether the immortal being was also named “Caesar” is irrelevant. For Jones' argument to work, therefore, she has to make the assumption that an action could remain the same one even if it were to have an utterly different target. But on this assumption, Brutus' action would have been the same event token even if he had stabbed a wooden statue of Caesar instead of Caesar himself. After all, a wooden statue of Caesar is no more different from Caesar himself than an immortal being is. The problem is that I have difficulties seeing how Brutus' assassination attempt could have been identical to a statue-damaging action. While it may be possible for an actual event to have occurred with different causes and effects, it is very unlikely that the causes and effects of an event could have been drastically different. Needless to say, whatever the causes and effects of Brutus' assassination attempt might have been, they would be radically unlike those possible events that could be the causes and effects of the statue-damaging action. This suggests that Jones' position is too extreme: Brutus' stabbing of Caesar couldn't have been a failed assassination attempt aimed at an immortal being named “Caesar”.

Where does Jones' position go wrong? My preferred analysis is that her position fails to appreciate the extent to which the properties of an event are constrained by the event's primary kind. It is impossible for an assassination attempt to have been a statue-damaging action because assassination attempts and statue-damaging actions fall under fairly different primary kinds. In light of our observations, I suggest that we use the following criteria to determine how different an event could have been:

(Contingent Properties of Events)

Let  $e$  be an event whose spatiotemporal location is  $L$ .

- (1) For any event  $f$ , we say that  $e$  *could have been*  $f$  if and only if:
  - (a)  $e$  and  $f$  are located in different possible worlds;
  - (b) the spatiotemporal location of  $f$  is also  $L$  (or the location corresponding to  $L$  in the world  $f$  belongs to);
  - (c)  $e$  and  $f$  fall under the same primary kind;
  - (d) for any property  $P$ , if  $e$  exemplifies  $P$  in virtue of falling under its primary kind, then  $f$  also exemplifies  $P$ .
- (2) For any property  $P^*$ , we say that  $P^*$  is a contingent property of  $e$  if and only if there is an event  $g$  such that  $e$  could have been  $g$  but  $g$  does not exemplify  $P^*$ .

Consider again my claim that some worlds contain **no-confidence** but not **speech** and **vote**.

According to my proposal, this is perfectly intelligible insofar as in those worlds **no-**

**confidence** is constituted by events whose primary kinds are identical to those of **speech**

and **vote**. Any event that is so constituted is an event that **no-confidence** could have been;

therefore, in any world where an event constituted this way occupies the location that is

occupied by **no-confidence** in the actual world, the event can be plausibly identified as **no-**

**confidence** in that world. If this is correct, we should reject the view that all properties

instantiated by an event are instantiated necessarily.

### 3.7 Conclusion

I have tried to show that the existing versions of hylomorphism about events suffer from several difficulties, including the failure to respond to the CP problem. I have also suggested that the CP problem can be solved if we think of constitution as the satisfaction of immediate and specific dependence needs. These complete the metaphysics part of my dissertation. Due to the orientation of this chapter, however, I had to set aside a number of issues deserving detailed analysis. I shall mention one of them here. There is a debate about whether the constitution relation and the parthood relation can be understood in terms of each other. As Evgine points out, answers to this question range from the view that the two are the same relation to the view that they are mutually independent (2011, 212–13). Discussions of this issue have largely focused on the constitution relation between physical objects and quantities of matter, but we can also raise the issue when theorizing about events. Here the issue becomes even more complicated. As Eric Olson notes, thorny problems arise when we try to relate temporal parts to the more ordinary kind of parts, i.e. the kind of parts whose possession is relativized to times (2006, 739). Arguably, though material constitution is more closely related to the ordinary kind of parthood, event constitution is more intimately bound up with temporal parthood. If so, while accounts of the relation between parthood and material constitution may shed light on the relation between parthood and event constitution, the latter issue requires additional analysis and is therefore worth pursuing.



## CHAPTER 4

### PERCEPTION: A HUSSERLIAN REPRESENTATIONAL VIEW

I take the most important issue in the philosophy of perception to be the problem of perceptual intentionality: what is the relation between the perceiver and the world?

According to the representational view, the answer is representation: the perceiver is related to the world by entering a mental state capable of receiving information from the world.<sup>40</sup> In this chapter, I defend a Husserlian version of the representational view, which I call the Husserlian dual-component view (HDC). To situate HDC in the theoretical landscape of perception research in philosophy, I shall begin with a few words about the representational view in general.

Suppose Albrecht is a tall person. If I see Albrecht, my experience tells me that Albrecht is tall. How should we describe my experience? We could say that Albrecht is the *object* of my experience or that my experience *represents* Albrecht. The question then becomes: how does my experience do that? A possible answer is that my experience represents Albrecht by exemplifying a proposition, which can be expressed by the sentence “Albrecht is tall”. The proposition is said to be the *content* of my experience, and my experience represents Albrecht by means of its content.<sup>41</sup> This is a very rough sketch of a possible version of the representational view, but it gives some idea of how the representational view approaches

---

<sup>40</sup> The claim that perception is representational is intensely debated in contemporary philosophy of mind. I follow Campbell in calling the view that affirms the claim *the representational view* and its main rival *the relational view* (Campbell 2002, 116). There are various ways to draw the distinction between the two views, and I will examine the distinction in greater detail in section 4.4.

<sup>41</sup> I will say more about the content-object distinction in section 4.1.

the problem of perceptual intentionality. The problem is understood as one about perceptual contents: what is the relation between the contents of a perceptual experience and the objects of the experience? To answer this question, at least two issues have to be addressed. First, how are perceptual contents structured? Are perceptual experiences similar enough to sentences such that they also express propositions, or is their structure rather different from anything linguistic? Second, what do perceptual contents represent? Do they represent nothing but clusters of properties, or do they in some sense represent the bearers of properties as well? Proponents of the representational view differ in both the relative importance they assign to these two issues and the responses they come up with.

The goal of this chapter is to defend a Husserlian account capable of responding to both issues. The account is Husserlian, because it is inspired by Husserl's ideas. But the account is not Husserl's own; in fact, this chapter contains very little exegesis of Husserl's texts, even though it could certainly use more. Instead, I draw heavily on the work done by three contemporary phenomenologists: A. D. Smith (2008), Walter Hopp (2010, 2011) and Michael Madary (2016), who depart from Husserl from time to time to address contemporary concerns. In addition, my view about how objects are related to perception is much indebted to John Campbell (2014), one the most prominent opponents of the representational view. While I take the resulting view to remain recognizably Husserlian, I would be willing to accept the charge that there are inconsistencies between my view and Husserl's own. Roughly, HDC says that the representation of objects is the joint achievement of two kinds of contents: those representing what is strictly in view and those representing what is expected to come into view. Take a magic show in which the performer produces a flying pigeon from a piece of paper. Carl watched such a show and found the

experience bewildering. Why? The reason is that what was strictly seen by Carl, i.e. a pigeon figure drawn on paper, did not exhaust Carl's experience. His experience was also constituted by the expectation that the figure would remain two-dimensional. A perceptual conflict thus arose when the figure suddenly became a real pigeon, and Carl felt baffled even before he could tell what made the experience baffling. According to HDC, the fact that the contents of perception have such dual components explains how a figure on paper, or a flying pigeon, can be perceptually represented.

HDC is faced with strong contenders. Two other versions of the representational view are much more widely discussed than HDC: the views on which perceptual contents are, respectively, linguistically and pictorially structured. We might think that perception derives its representational capacity from the subject-predicate structure of its content, which mirrors the object-property structure of the things we encounter in perception. Alternatively, we might think that perception is representational because it is in some sense an internal image of the world. These versions of the representational view are intuitively appealing – arguably more so than HDC. But as I see it, the propositional and pictorial accounts of perception fail to appropriately describe the contents of perception. It is hard to both substantiate the claim that perceptual contents are sufficiently similar to propositions (or pictures) and explain away the dissimilarities between perceptual contents and propositions (or pictures). As a result, these relatively mainstream representational views bring with them presuppositions that make them untenable accounts of perception – or so I shall argue. However, the strongest challenge to HDC does not come from the other versions of the representational view. It comes from the main rival of the representational view: the relational view. According to the relational view, the very attempt to understand perception

in terms of representation is wrongheaded: any representational content an experience might have is derivative of a perceptual relation that is non-representational. The proponents of the relational view have mounted several objections against the representational view. I will argue that HDC is immune to them, and this makes HDC a promising version of the representational view.

This chapter is organized as follows. Section 4.1 examines the idea of representation in more detail and distinguishes between its different aspects. Once the distinctions are in place, I argue against the proposition- and picture-based representational views in section 4.2. HDC is then presented in section 4.3, which makes up the bulk of the chapter. Section 4.4 briefly considers the relational view: I examine two objections raised by Campbell against the representational view and consider how HDC might respond to them. After that, section 4.5 concludes the chapter.

## 4.1 Vehicles, Contents and Objects

As a preliminary to examining the representational view, let's consider three important aspects of representation: the vehicle, content, and object of an experience.<sup>42,43</sup> I follow Fred Dretske (1981) and characterize representation in terms of information, though my understanding of information differs substantially from Dretske's own. I take a mental

---

<sup>42</sup> Given my focus, I will use the terms "experience", "perception" and "perceptual experience" to refer to *conscious visual experiences had by human beings*, unless otherwise noted. There is no denying that experiences of other sensory modalities and cases of animal consciousness are also philosophically significant. However, those issues are complicated and cannot be appropriately addressed here.

<sup>43</sup> I borrow the tripartite distinction from Hopp (2011, chap. 1). For additional discussions of the content-object distinction, see Husserl ([1913a] 1970, 1:199–200), Crane (2006) and Burge (2010, 413).

event to be *representational* or *contentful* if it *carries information* about something other than itself. Such events are also said to have *representational contents*, or simply *contents*. For mental events to count as perceptual experiences, they typically have to carry information about entities that are not components of the perceiver. But what is it for a perceptual experience to carry information? Consider what Susanna Siegel says about *conveyed contents*. According to her, “a content is conveyed by experience if it would be a content of explicit beliefs that are natural to form on the basis of visual experience”, or “if it enables the experience to guide bodily actions”, or “if it is manifest to introspection that it is a content of experience” (2010, 51). On my view, a perceptual experience carries information if, roughly speaking, it bears conveyed contents in Siegel’s sense. That is, a perceptual experience carries information if it provides inputs to the perceiver’s other mental processes, such as beliefs, actions and cases of introspection. There are nevertheless two reasons that I shall not simply adopt Siegel’s criteria of conveyed contents. First, as Siegel herself notes, some think that perceptual experiences never have the same contents as beliefs (2010, 51 fn. 21). These people hold that Siegel’s first criterion is never satisfied by any experience, and I happen to be one of these people. Second, more generally, it is hard to figure out how the three criteria are related to each other. There may be something special about beliefs, actions and introspection, but Siegel doesn’t tell us what it is that makes them special. Why exactly must an experience be related to these states for it to bear conveyed contents? Without an answer to this question, the very idea of conveyed contents remains opaque.

As an alternative, I propose the following criterion of information carrying: an experience carries information about an object if, possibly, the experience enables one to better carry out a bodily or mental task that involves the object. That is, the reason an

experience is said to carry information about the perceiver's environment is that it enables the perceiver to interact with her environment in ways that are otherwise impossible. For example, suppose Edmund was shown an image for one second and then asked to determine the number of rectangles in the image. Edmund succeeded, but he obviously would have failed had he not seen the image. His visual experience thus enabled him to better carry out the task of counting rectangles. We can then say that Edmund's experience carried information about, or represented, the rectangles in the image. As another example, suppose Franz is a professional athlete that specializes in hurdling. He jumps over a hurdle whenever he sees one, but he simply cannot do that if he doesn't undergo the right kind of visual experiences before jumping. His visual experiences thus enable him to better carry out the task of jumping over hurdles. On the present view, his experiences carry information about hurdles. Note that an experience can carry information about an object even if one doesn't actually perform any task that involves the object. Insofar as it is possible to better perform a task with the help of an experience, the experience represents the objects involved. Take the experience I undergo when I look at the keyboard of my laptop. I am not typing, but if I were to type, I would know where to put my hands because of the experience. My experience thus carries information about the keyboard. Also note that I take this criterion to capture a sufficient condition for information carrying, but it doesn't specify any necessary conditions. The criterion is compatible with the possibility that experiences are capable of carrying information in a different way, such as in the way described by Dretske.

Why bother with an alternative criterion of information carrying instead of just using Dretske's own? The reason is that the alternative criterion makes the intimate relation between consciousness and representation more salient. One can barely perform any mental

or bodily task without some form of consciousness. In contrast, given Dretske's naturalistic project, consciousness is conspicuously absent in his definition of information carrying.

Dretske defines *informational content* in terms of conditional probability: "A signal  $r$  carries the information that  $s$  is  $F$  = The conditional probability of  $s$ 's being  $F$ , given  $r$  (and  $k$ ), is 1 (but, given  $k$  alone, less than 1)" (1981, 65). Here  $s$  is the source of information and  $k$  the background knowledge one has about the source (1981, 80–81). Suppose Franz's fridge always starts beeping 30 seconds after its door is opened. Franz knows this, and he hears the beep now. In this case, the conditional probability of the fridge door's being open is 1, given Franz's background knowledge about the fridge and the occurrence of his auditory experience. By Dretske's definition, Franz's auditory experience carries the information that the fridge door is open.

Dretske's definition is precise, but it is unclear to me that it is suitable for the purpose of describing conscious experiences. Suppose I am looking at my laptop. One could build a complicated signal detector that enters a particular state only if its position relative to my laptop is the same as my present position relative to my laptop. Given the machine's state, the conditional probability of my laptop's presence is 1, so the machine's state carries information about my laptop. When applied to the machine, this account of information makes perfect sense. But I doubt that my present experience carries information about my laptop in precisely the same way the machine's state does – arguably the machine doesn't exemplify any form of consciousness. It is thus unclear how consciousness fits into Dretske's view about contents.

To be sure, Dretske does have something to say about consciousness. He adopts the widely held distinction between creature and state consciousness and defines the latter in

terms of the former. Creature consciousness is attributed to “beings who can lose and regain consciousness and be conscious of things and that things are so” (1995, 98), so it serves as a mark that distinguishes entities with certain mental capacities from those without them. In contrast, state consciousness is attributed to a type of states conscious creatures are in, and such states are “the ones that make you conscious of things” (1995, 116). That is, a state is conscious if and only if it underlies some form of creature consciousness. To accommodate Siegel’s conveyed contents in Dretske’s framework, we could take a conscious experience to consist of two components: the first component carries information about the experience’s objects and the second component affords access to the information carried by the first component. Call the first the *representational* component and the second the *experiential* component. Now, since the experiential component merely functions to make information consciously available, it has no representational capacity. It follows that whatever contents an experience has, they are entirely derived from the representational component of the experience. If so, there is no need to invoke consciousness when we define contents. My concern about Dretske’s view is misguided.

Such a response on behalf of Dretske’s view assumes that we can successfully differentiate between the representational and experiential components of an experience. We can only do this if consciousness merely makes hitherto unavailable information available, without changing the information at all; that is, we can only do this if consciousness contributes nothing to representation. It is unclear that this assumption is warranted. Even Dretske himself seems to hold that conscious states carry different information than unconscious ones: “The function of sense experience, the reason animals are conscious of objects and their properties is to enable them to do all those things that those who do not



have it cannot do” (1995, 121). For conscious states to have this function, it seems that their capacity to bring about creature consciousness must contribute to their information-carrying capacity. Otherwise, the animals that are conscious of objects and their properties would not be able to do anything that cannot be done by the animals that are not conscious of objects and properties. It is therefore premature to assume that perception has a representational component and an experiential component that are independent of each other. In any case, whether perceptual consciousness can be explained in terms of the information-carrying capacity of perception is a heatedly debated issue,<sup>44</sup> and I prefer a view about contents that remains neutral about the issue. I will therefore stick to the alternative criterion of information carrying given earlier.

Representational contents must be distinguished from represented objects. On my construal, whatever possibly exists is an *entity*, and an entity about which an experience carries information is an *object* of the experience. When I see a pumpkin, the feature of my experience that carries information about the pumpkin is the content of the experience, whereas the pumpkin itself is the object of the experience. So construed, *being an object* is a relational property an entity has in virtue of being represented, so there are no entities that are intrinsically objects. It follows that objects must be distinguished from *individuals* or *substances* like persons and mosques. The class of objects includes more than individuals: such entities as properties, events and quantities of matter can also be objects, so long as there are experiences that bear the representational relation to them.

---

<sup>44</sup> For a discussion of some physical relations that often figure in reductionistic explanations of mental contents, see Pautz (2009, sec. 4).

I will not spend much time defending the content-object distinction, because I think Hopp (2011) has already made a very strong case that the distinction ought to be acknowledged.<sup>45</sup> However, I would like to stress the relevance of the distinction to the issue of whether perceptual contents should be thought of as propositional, pictorial or neither. In his recent defense of a pictorial view, Mohan Matthen denies that the distinction between depictions and propositions is one between two kinds of contents: “The correct contrast is between *sentential vehicles* of meaning and *imagistic vehicles*. Both sentences and images express propositions, but they do so in different ways” (2014, 268). I disagree, and I believe the main problem of Matthen’s claim is that it lumps together the content and object of a representation. Though a detailed description of, say, Rembrandt’s *The Night Watch* might convey roughly the same information as the painting itself, this is not because they have the same content. The content of the description is determined partly by the meanings of the lexical items in the description and partly by the rules of deriving truth conditions from those meanings. However, whatever rules we are going to use to derive truth conditions from the meanings carried by the basic elements of the painting, they will differ from the rules applicable to any description of the painting. If so, the reason that the painting and the detailed description of it convey similar information is that they represent the same objects, not that they have the same propositional content. The painting and the description both represent the militia guards themselves, for example.

But isn’t the content-object distinction just the vehicle-content distinction? Isn’t the difference merely terminological? The answer is “no”, and it is now time to make explicit the

---

<sup>45</sup> See footnote 43.

reason that we need a tripartite distinction between vehicles, contents and objects. The reason is nicely illustrated by how semantics is typically done in the Montagovian tradition.<sup>46</sup> Take the sentence “the London Eye rotates slowly”. At the highest level of syntactic decomposition, the sentence has two components: “the London Eye”, a noun phrase, and “rotates slowly”, a verb phrase. Given the sentence’s syntax, the sentence’s meaning also has two components at the highest level of semantic decomposition: the meaning of “the London Eye” and that of “rotates slowly”. The sentence needs to be so decomposed semantically *because of* the way it is decomposed syntactically. Now, the meaning-components also have to be understood in a twofold way, in terms of Fregean *Sinn* and *Bedeutung*. The extension of “the London Eye” is an individual, the London Eye, while the extension of “rotates slowly” is a function from individuals to truth values. The function yields truth if and only if the individual it takes rotates slowly. The extensions are not intensions. The intension of “the London Eye” is a constant function that takes any world-time pair and yields the London Eye itself. In contrast, the intension of “rotates slowly” is a function that specifies, for any world-time pair, a set of individuals that rotate slowly. On this picture, there is a clear distinction between vehicles, contents and objects: it is just the distinction between written symbols (or uttered sounds), intensions and extensions.

If we understand semantics this way, Matthen’s claim is problematic. To say that a picture expresses a proposition is to say that the picture can be decomposed into two syntactic parts. The first part carries a meaning that denotes an individual, whereas the second part carries a meaning that denotes a function from individuals to truth values. I

---

<sup>46</sup> See, for example, Dowty et al. (1981).

don't see how this way of decomposing a picture makes sense. Which syntactic part of *The Night Watch* denotes nothing but an individual? Since the syntax of pictures is radically different from that of sentences, pictures and sentences cannot have the same semantics. Rather, they differ with respect to both vehicles and contents: whereas imagistic vehicles carry depictive contents, sentential vehicles carry propositional contents. So construed, the vehicle-content distinction is obviously different from the content-object distinction, and both must be acknowledged. One doesn't have to understand semantics in the Montagovian way, of course. Nevertheless, even if one adopts alternative accounts of syntactic expressions, intensions and extensions, the distinction itself is still very compelling. I will therefore assume the tripartite distinction throughout this chapter.

Now that we have examined how contents are related to vehicles and objects, we can move on to some representative versions of the representational view. In particular, I will examine two views, which take representational contents to be propositional and pictorial, respectively. An analysis of the problematic features of these two views will set the stage for our discussion of a Husserlian view.

## 4.2 Two Common Views

### 4.2.1 Propositionalism

Many influential theorists take the contents of perceptual experiences to be propositional. John McDowell offers the following analysis of a veridical experience: "*That things are thus and so* is the content of the experience, and it can also be the content of a judgment: it becomes the content of a judgment if the subject decides to take the experience

at face value. So it is conceptual content” (1996, 26). Alex Byrne also invokes propositions in defending intentionalism: “the propositional content of perceptual experiences in a particular modality (for example, vision) *determines* their phenomenal character” (2001, 204). His conception of a proposition is that of “an abstract object that is a truth-bearer, that is the object of some propositional attitude-like psychological states, and that determines a possible-worlds truth condition” (2001, 201 fn. 5). Susanna Siegel takes a similar approach: “there are a number of abstract objects that contents could be, corresponding to different kinds of propositions, and we can ask which such abstract objects are best for characterizing the contents of experience” (2010, 77). Following Madary, I will call such a view *perceptual propositionalism* (Madary 2016, 65).

It is not easy to spell out what exactly perceptual propositionalism entails. Arguably, what a philosopher means by the term “proposition” has to be determined within the context of her theory, so there might not be a common core of the various views about propositions. Despite this, I think Byrne’s view is relatively uncontroversial.<sup>47</sup> I shall therefore take propositionalism to be the view that perceptual contents are at least capable of assuming the roles Byrne attributes to propositions. Though this version of propositionalism is still quite abstract, it is by no means free of theoretical baggage. Consider Gareth Evans’ famous Generality Constraint:

---

<sup>47</sup> Though Byrne’s view may be relatively uncontroversial, it is certainly not universally accepted. The term “object” in Byrne’s characterization of propositions is ambiguous. Consider my belief that Mount Fuji is in Japan. Which entity is the object of my belief, Mount Fuji or the proposition that Mount Fuji is in Japan? As I made clear in section 2.3, I would only accept the first answer. But this is not the place to press the issue.

...we cannot avoid thinking of a thought about an individual object  $x$ , to the effect that it is  $F$ , as the exercise of two separable capacities; one being the capacity to think of  $x$ , which could be equally exercised in thoughts about  $x$  to the effect that it is  $G$  or  $H$ ; and the other being a conception of what it is to be  $F$ , which could be equally exercised in thoughts about other individuals, to the effect that they are  $F$ . (Evans 1982, 75)

Take, for example, the belief that Germany is a member of the European Union and the belief that Germany is the country in which BMW is based. There is an important similarity between the beliefs: in both, one has to think about Germany. Since propositions are supposed to be what are believed when one acquires beliefs, the similarity must be specifiable in terms of propositions. There is also an important similarity between the belief that Germany is a member of the European Union and the belief that Austria is a member of the European Union: they both require a conception of what it takes to be a member of the European Union. Again, the similarity must be specifiable in terms of propositions. It is hard to see how the similarities just mentioned can be accommodated without assuming that propositions feature some kind of compositionality. From the Generality Constraint, therefore, it is a short step to the compositionality of propositions. In fact, we probably have to say that the compositional principles of propositions are those specified by first-order logic; otherwise, we would soon run into pairs of possible beliefs whose similarities are inexplicable.

Such theoretical baggage makes perceptual propositionalism problematic. I take the main issue with propositionalism to be that it decomposes perceptual contents in the wrong way. A proposition consists of at least a constant and a predicate. Propositionalism thus entails that there are perceptual constants and perceptual predicates. In addition, it should be possible for any constant to combine with any predicate – this possibility goes to the very

heart of the standard semantics of propositions. It follows that the perceptual propositionalist has to take any combination of a perceptual constant and a perceptual predicate to be well-formed. That has unfortunate consequences, as Hopp has pointed out in arguing against conceptualism about perceptual contents. The propositionalist has to accept any well-formed “perceptual proposition” as a possible perceptual content, whether or not there is any experience capable of exemplifying the content. To use an example of Hopp’s, the sentence “The Hammerklavier Sonata is tall” is well-formed, but it is not a content that can be exemplified by any experience (2011, 128). This is the case even though the Hammerklavier Sonata and the property of being tall are both perceptible. And additional examples abound. For instance, insofar as I focus on the Hermann grid, I cannot fail to see the grid as containing grey blobs. The impossibility has nothing to do with the physical features of the Hermann grid; the grid contains no grey blobs. Rather, the impossibility results from the very way human perception works. Given the nature of human perception, a perceptual representation of the Hermann grid simply cannot be combined with a perceptual representation of the property *containing uniformly white lines*. According to propositionalism, however, that shouldn’t be impossible: “the Hermann grid contains uniformly white lines” is certainly well-formed.

The unrestricted combination of perceptual constants and predicates is not the only problem facing propositionalism. In the standard semantics, a complex proposition is always equivalent to a conjunction of multiple atomic propositions. We have the former if and only if we have the latter. Take two experiences whose contents are expressed by the following sentences, respectively: “the London Eye is white” and “the London Eye is rotating”. Suppose the experiences occur successively without any gap. Propositionalism entails that

they have a fusion whose content is expressed by “the London Eye is white and rotating”. There is nothing surprising about this complex experience; it’s just the temporally extended whole formed by the successive experiences. So far, so good. But now consider the Necker cube diagram. Suppose I just began to see the cube as oriented in one of the two ways less than 500 milliseconds ago; it will be difficult for me to immediately “switch back” and see the cube as oriented in the other way. The difficulty has nothing to do with the physical features of the Necker cube diagram – the diagram itself is not even three-dimensional. The difficulty has its origin in human perception: what perceptual representations can be combined at a given moment partly depends on what representations were combined at the preceding moment. But propositionalism is unable to accommodate this. On propositionalism, undergoing an experience whose content can be expressed as “the Necker cube is oriented in direction  $D_1$  at  $t_1$ ” shouldn’t make it any less likely for one to undergo an experience whose content can be expressed as “the Necker cube is oriented in direction  $D_2$  at  $t_2$ ”. After all, if we combine the contents of the two experiences into a complex proposition, the resulting proposition doesn’t contain any kind of contradiction. As reflection on everyday life shows, however, what one experiences at a given moment puts substantial constraints on what one can experience at the next moment. Propositionalism easily gives rise to a discrete conception of experience, which gets perception wrong.

Furthermore, there are many logical or inferential relations that may hold between propositions but not between perceptual contents. Note that I do not mean to suggest something as strong as Tim Crane’s view. He outright denies the existence of any logical relations between experiences: “Pictures do not imply one another; they cannot be negated or disjoined. *In this way they are like experiences*” (2011, 89; my italics). In my opinion, this is too



strong. Suppose my experience exemplifies content  $c$ , which represents a pumpkin in front of me. *In virtue of*  $c$ , my experience exemplifies content  $d$ , which represents the spatial region in front of me as occupied by an object. Can't we say that  $c$  implies  $d$ ? On the one hand, the relation between  $c$  and  $d$  is not contingent. It is necessary that  $d$  is instantiated if  $c$  is instantiated. On the other hand,  $c$  and  $d$  are both abstract entities. The fact that they always go hand-in-hand cannot be explained in terms of such relations as causation. It then emerges that the relation between  $c$  and  $d$  is very much like the entailment relation between propositions. To be sure,  $c$  and  $d$  are not propositions. But that shouldn't prevent us from acknowledging how much the relation between  $c$  and  $d$  resembles entailment. As the impressive work of Jan Westerhoff (2005) shows, we can even define a logic that applies to depictive contents.

That said, many logical relations still fail to hold between perceptual contents. This is not because perceptual contents never entail each other; they sometimes do. Rather, it is because any relation that may be called "perceptual entailment" is different from the relation of logical entailment. The rule of universal generalization guarantees that certain propositions about individuals entail universally quantified propositions, but the same pattern cannot be found in perception. Similarly, a contradictory proposition entails every other proposition, but no perceptual content entails every other perceptual content. The list goes on. By taking perceptual contents to be propositions, therefore, one is forced to attribute a set of laws to perceptual contents that are incompatible with perceptual contents. It is worth emphasizing that these consequences are not those of a very robust conception of propositions. Rather, from the idea that propositions are what we believe, it is a short step to the conclusion that perceptual propositionalism has the above consequences. We arrive

where we are simply by considering the range of possible beliefs and the compositionality of propositions needed to accommodate the variety of beliefs. The best response, I submit, is to abandon perceptual propositionalism. The cost of holding onto such a view is simply too high.

#### 4.2.2 *Pictorialism*

For those dissatisfied with propositionalism, a natural alternative is the view that perceptual contents are in some sense pictorial, imagistic or depictive. Such a view is appealing especially in light of the similarities between vision and visual imagery. The psychologist Stephen Kosslyn has been working on mental imagery for decades. He says, in a recent co-authored book, “A fundamental assumption of the present approach is that visual imagery evokes many of the same processing mechanisms used in visual perception” (Kosslyn, Thompson, and Ganis 2006, 135). Such a view is not without proponents in philosophy. Commenting on Kosslyn’s earlier work, Michael Tye says, “Like Kosslyn, I believe that there is a medium shared by imagery and vision” (1991, 91). Similarly, Jerry Fodor argues that “it is (empirically) plausible that some perceptual representation is iconic” (2007, 107). These remarks suggest that images provide a fruitful way to understand vision.

To say that vision shares a medium or processing mechanism with mental imagery is not yet to say that visual experiences have depictive contents, of course. It could be that the contents of visual experiences are still propositional; it’s just that the biological states that are the vehicles of such contents are depictive. But on the basis of the claim that the visual system makes use of depictive representations, it is tempting to make the case that visual experiences have depictive contents. After all, what representations the visual system

employs puts substantial constraints on what contents visual experiences have. There are indeed philosophers who explicitly endorse a more-or-less depictive account of perceptual contents, even though the reasons they give are diverse. According to Crane, “a view on which the content of perceptual experience is more like the content of a picture gives a better account of the fact that experiences can be accurate or inaccurate than the propositional-attitude theory does” (2011, 93). According to Mohan Matthen, “Perceptual experience has image content”, and this is because “Perceptual experience presents the subject with a spatiotemporally ordered array of sensory qualities” (2014, 266). In what follows, I will call the view that perceptual contents are depictive *perceptual pictorialism*.

Just as in the case of perceptual propositionalism, it is not entirely clear whether there is a common core shared by all versions of perceptual pictorialism. The closest I can find is a thesis espoused by Fodor, which appears in somewhat similar forms in several accounts. Fodor calls the thesis the *Picture Principle*: “If P is a picture of X, then parts of P are pictures of parts of X” (2007, 108). In other words, a distinctive feature of depictive representations is that the meanings of their parts survive arbitrary decomposition. Consider the sentence “the outermost arc of a rainbow is red”. If one divides the sentence into “the outermost arc of” and “a rainbow is red”, neither can be regarded as a partial representation of the state of affairs described by the original sentence. In contrast, depictive representations can be arbitrarily decomposed. If I cut a painting of a rainbow into pieces, a red piece still represents a part of the outermost arc of the rainbow. Kosslyn proposes a similar semantics for depictions: “each part of the depiction must correspond to a visible part of the object such that the distance among the points on the object are preserved by the corresponding ‘distances’... among their representations” (1984, 107). This is a variant of

the picture principle, though the preservation requirement does put some additional constraints on decomposition. Tye interprets the preservation requirement as saying that depictions must represent greater apparent distance by a greater number of representation-

parts. If object  $a$  appears to be located farther from  $b$  than from  $c$  in my experience, then my experience must have more parts representing



**Figure 1. The London Eye**

Adapted from Milas Bowman, *London Eye*. [CC BY-SA 2.0, <https://creativecommons.org/licenses/by-sa/2.0/>], via Wikimedia Commons.

the region connecting  $a$  and  $b$  than parts representing the region connecting  $a$  and  $c$  (Tye 1991, 36). If this requirement is met, nevertheless, Kosslyn and Tye seem to allow all kinds of decomposition. Suppose I am looking at the London Eye. It is not rotating now, and the way it looks to me is captured by Figure 1. Kosslyn and Tye seem to accept that the representation of region  $R$  is a part of the content of my experience *in the same sense* as the representation of car  $A$  is a part of the content. If so, their version of the picture principle is still very liberal with regard to decomposition.

I find this analysis very puzzling, because the representation of region  $R$  is not a meaningful unit of the content of my experience – at least not in the sense that the representation of car  $A$  is. If I undergo a series of experiences of the London Eye, it is quite likely that I will be able to introspectively identify a representation of  $A$  in all these experiences. Moreover, while these representations would not be qualitatively identical to each other, their introspectively accessible features will be salient enough that they can be readily distinguished from the representations of, say, car  $Z$  or hub  $O$ . I doubt that the same

can be said of my representation of region  $R$ . It would seem that the moment I blink, I would no longer be able to perceptually identify region  $R$  or introspectively identify the experiential representation thereof. Therefore, to produce an adequate description of my experience, we need to account for the difference between the experiences of  $A$  and  $R$ . According to the Picture Principle, however, there simply cannot be any structural difference between them.

In fact, I am not even sure that the Picture Principle is applicable to physical pictures. Elaborating on the principle, Fodor says, “an icon is a homogeneous kind of symbol from both the syntactic and the semantic point of view. Each of its parts *ipso facto* gets a semantic interpretation according to the same rule of interpretation that applies to each of the others (viz. according to the Picture Principle)” (2007, 108–9). Hence, while in dividing up a sentence one needs to make sure that a word does not get cut in half, there are no pictorial parts that must be grouped together for them to play their representational roles. Every pictorial part plays the same representational role: to picture the part of the object to which it corresponds. Fodor’s claim seems quite untenable to me; there is an abundance of counterexamples he simply fails to take into consideration. Suppose I cut up a drawing of the brain according to brain areas. Then I find the parts corresponding to the parietal and occipital lobes and cut them into two piles of tiny pieces. According to the Picture Principle, if I pick up one piece from each pile, the two pieces of drawing represent different parts of the brain. This is the case even if the pieces are barely distinguishable from each other. Such an analysis strikes me as problematic; it seems that the two pieces lose their representational functions once they stop being parts of a pictorial system. If so, however, the relation a pictorial part bears to another may very well affect what the pictorial part represents. As a

matter of fact, there are all kinds of ways a pictorial part may be related to another part. Given that only some of the relations can be preserved by a certain manner of decomposition, a manner of decomposition is acceptable only if it preserves the syntactically and semantically relevant relations in a picture. Fodor is therefore wrong to claim that pictures can be arbitrarily decomposed.

Additional considerations against the Picture Principle can be found in Westerhoff's account of pictures. He distinguishes between different kinds of parthood relations that can be found in pictures, and two of them are subpictures and (mere) parts (2005, 607–9). Consider Rembrandt's *The Night Watch* again. A part of the painting's content that represents a militia guard is a subpicture, while the part of the content that corresponds to a diagonal of the painting is just a mere part. One of the reasons for drawing the distinction is that "collections of pixels from a picture can fail to be either a part [viz. a mere part] or a subpicture" (2005, 609). For example, one could cut out a cross from the black background in *The Night Watch*. The cross would seem to be a picture, but it would not have been one in the original context. We would not be able to make sense of this phenomenon if there were no distinctions between kinds of parts. Westerhoff's analysis again shows that the Picture Principle is false. If subpictures and mere parts are different from each other, decomposition cannot proceed in an arbitrary manner. One would never get subpictures if one decomposes a picture pixel by pixel.

If the reason to take perceptual pictorialism seriously in the first place is that we might be able to better understand perception by considering some of the important characteristics of physical pictures, then the Picture Principle is not helpful in this regard. The feature it attributes to physical pictures fails to do justice to their nature. To be sure,

proponents of perceptual pictorialism don't have to be committed to the Picture Principle. Something like Westerhoff's logic of pictures provides a much more solid foundation of pictorialism. However, it remains to be seen whether depiction and perception have enough in common to warrant a pictorial account of perception. For now, I will set pictorialism aside.

### 4.3 The Husserlian Dual-Component View

#### 4.3.1 *Husserl on Perception*

The goal of section 4.3 is to defend a Husserlian alternative to both perceptual propositionalism and pictorialism – the Husserlian dual-component view (HDC). Obviously, the central ideas in HDC go back to Husserl. I will therefore offer a brief overview of the Husserlian theoretical background before I formulate HDC in 4.3.2. Husserl's view on perception was gradually developed throughout his career, so there is no single work that contains everything Husserl has to say about perception. However, the main ideas I am concerned with in this chapter can be found in Husserl ([1913a] 1970, [1913b] 1970, [1966] 2001). Perception assumes a central role in Husserl's theory of knowledge. Knowledge consists minimally of verified thoughts, so Husserl asks how thoughts can be verified. The mental states that verify thoughts are called *intuitions*. As Husserl puts it, thoughts are “‘illustrated’, or perhaps ‘confirmed’ or ‘fulfilled’... or rendered ‘evident’” by intuitions ([1913a] 1970, 1:174). Perceptual experiences are intuitions of the most fundamental kind, because they play the role of confirming thoughts most successfully. Husserl describes perceptual experiences as *self-giving*, which means “that every perception within itself is not

only... a consciousness of its object, but that it gives its object to consciousness in a distinctive manner. Perception is that mode of consciousness that sees and has its object itself in the flesh” ([1966] 2001, 140). When perception works properly, it brings into view the objects we think about. Undergoing perceptual experiences of the appropriate kind thus puts one in a highly favorable position to determine whether one’s thoughts are true. If my friend and I disagree about whether Mount Fuji is in Japan, the way to settle the issue at the most fundamental level is to go to Japan and see for ourselves whether Mount Fuji is there. Notably, as the passage just quoted shows, Husserl thinks that perception is able to do so partly because of its distinctive first-person features. What it is like to undergo a perceptual experience is radically different from what it is like to entertain a thought, and this is supposed to explain, at least in part, why the epistemic functions of thoughts and perceptual experiences are so different.<sup>48</sup>

We have seen that Husserl stresses the fact that perception brings objects into view. The idea should nevertheless not be interpreted as saying that the object of one’s experience is exhausted by what is clearly in one’s view. A distinctive feature of Husserl’s account of perception is that he takes expectations to be the very components of conscious experiences themselves. Husserl frequently uses the term “horizon” to describe expectations. A horizon represents what appears, as it were, on the horizon; it represents something that could be,

---

<sup>48</sup> It should be noted that despite the importance assigned to perception, the class of intuitions includes more than perceptual experiences: “The essential homogeneity of the function of fulfilment... obliges us to give... to each fulfilling act whatever the name of an ‘intuition?’...” (Husserl [1913b] 1970, 2:280). For example, memory is not perception, but it is often possible to justify one’s beliefs on the basis of memory. If I remember that I just bought some bananas five minutes ago, it is perfectly reasonable to believe that there are bananas in my backpack. Cases like this suggest that certain forms of memory should also count as intuitions.



but has not yet been, clearly seen. According to Husserl, a horizon “prescribes a rule for the transition to new actualizing appearances” ([1966] 2001, 42). To illustrate the idea, suppose I am looking at my car. While its windshield is visible from where I stand, its right wing mirror is obstructed from view by a parking meter. Despite this, my experience doesn’t represent my car as missing its wing mirror. If someone asks me what I see, my answer would be “a car whose right wing mirror is obstructed from view”, not “a car missing its right wing mirror”. The Husserlian explanation of this is that I expect to see the wing mirror when I move. Since the wing mirror is represented in my expectation and my expectation is a component of my experience, the car is not represented as missing the wing mirror.

Horizons play a unifying role in perception. In the car example, my expectation indicates to me that the currently experienced appearance of a parking meter will be gradually replaced by the appearance of a wing mirror. Such an expectation explains why my experiences can cohere or conflict with each other. As Jeffrey Yoshimi summarizes Husserl’s view, “When things go as expected, our current experiences ‘fulfill’ (*erfüllen*) our previous expectations... Husserl also speaks of this in terms of ‘harmonious’ (*ein stimmig*) perceptual processes. When things do not go as expected, our experiences ‘frustrate’ (*enttäuschen*) our previous ‘expectations’ (*Erwartungen*) or ‘intentions’ (*Intentionen*)...” (2009, 124). It has to be noted that Husserl’s notion of a horizon is quite broad. As Smith points out, the horizon of an experience has many components, and one of them is one’s present awareness of the past experiences that led to the present experience (2008, 324). Given that horizons are supposed to be features of conscious perception, Husserl’s view entails that perceptual contents are much richer than most contemporary philosophers would take them to be.

Husserl's ideas can be developed in several directions, depending on how one chooses to answer certain key questions. Philosophers working in the phenomenological tradition have developed several versions of the representational view on the basis of Husserl's ideas, and I will pay particular attention to their responses to two questions. First, granted that the content of an experience has an expectational component, how is the expectational component related to the other components? Second, on what conditions do two experiences have the same content? In particular, can two experiences that represent different objects have the same content? As mentioned earlier, I will focus on the work done by Smith, Hopp and Madary. However, my preferred responses to the issues aren't quite the same as theirs. On the one hand, Madary appears to think that the content of an experience is completely derived from its expectational component. I disagree with his view and follow Hopp in holding that the content of an experience is not exhausted by its expectational component. On the other hand, Smith and Hopp argue that experiences of different objects cannot have the same content. I am inclined to think that this is not the case. These concerns will be spelled out in due course.

#### 4.3.2 *HDC: A First Formulation*

Before I describe my departure from the views I draw on, I will first formulate a version of HDC by selecting and modifying elements of those views. In the sections that follow, I will justify my theoretical choices by raising some objections to the views. Our point of departure here is a distinction drawn by both Hopp and Madary, which divides perceptual contents into two kinds. Hopp holds that perceptual contents include both *intuitive* and *horizontal* ones (2011, 130). Similarly, Madary speaks of *factual* and *AF* contents,

where the name of the latter comes from the Husserlian ideas of anticipation and fulfillment (2016, 59–60). I will adopt Madary’s term of “factual content” for contents of the first kind. Since I prefer a term that wears its meaning on its sleeve, I will use the term “expectational content” for contents of the second kind. The two kinds of contents are posited to capture the Husserlian ideas mentioned in the previous section. When one sees an object, the factual contents of one’s experience represent the visible properties of the object, while the expectational contents represent the hidden properties. In the car example mentioned earlier, my representation of the windshield is the factual content of my experience, while my representation of the right wing mirror is the expectational content of my experience.

As I conceive of it, HDC is just the view that the content of a perceptual experience has these two components:

(The Husserlian Dual-Component View: A First Pass)

The content of a perceptual experience consists of factual and expectational contents.<sup>49</sup>

However, though the two kinds of contents have been characterized intuitively, they need to be more precisely defined. Let’s begin with factual contents. According to Madary, to say that a perceptual experience has factual contents is to say that it “represents factual properties” (2016, 44). *Factual properties* are distinguished from *perspectival properties*; while the former “are properties that can, in principle, be perceived from many perspectives”, the latter “are properties that can only be perceived from a particular perspective” (2016, 28).

---

<sup>49</sup> My official formulation of HDC will be given in section 4.3.5.

Consider a corner of a rectangular figure I am looking at. While its property of being acute-angled is perspectival, its property of being right-angled is factual.

It is not enough to say that factual contents are representations of factual properties; we want to know more about the representations themselves. How are they individuated? The same object can obviously look different to one from different perspectives. Therefore, one's spatial location must be taken into consideration when determining the factual contents of one's experience. But this is not enough. We should also consult Madary's insight that visual attention is significant in our attempt to understand visual experiences. Madary notes that one's representations of the objects at the periphery of one's visual field are very indeterminate. To illustrate this point, he cites a study by Freeman and Simoncelli (2011). The researchers created a pair of images and instructed the subjects to fixate on the center of each image. The periphery of one of the images was distorted, and the researchers aimed to measure how much distortion would make the two images appear different. Madary stresses the fact that "the periphery can be distorted quite a bit without losing the indistinguishability between the images" (2016, 37). The study clearly shows that one's spatial location is far from sufficient to determine how an object looks to one; the same object can appear radically different before and after a saccade. It follows that the contents of experiences must be individuated in a very fine-grained way; there is no reason to assume that the content of one's experience would remain identical after a saccade.

My definition of factual contents respects Madary's observations. It consists of two clauses:

(Factual Contents)

Let  $e$  be an experience that whose subject is  $s$  and whose object is  $o$ . Then  $e$  has a factual content if and only if  $e$  has some content  $c$  such that:

- (1)  $c$  represents a factual property of  $o$  that is visible to  $s$ ;
- (2) for any experience  $e^*$  whose subject is  $s^*$  and whose object is  $o^*$ ,  $e^*$  has content  $c$  if and only if EITHER:
  - (a) given the spatial location and visual fixation of  $s^*$ ,  $o^*$  appears to  $s^*$  in exactly the same way  $o$  appears to  $s$ ; OR
  - (b)  $e^*$  is a nonveridical experience subjectively indistinguishable from  $e$ .<sup>50</sup>

Condition (1) is straightforward. Take a slice of a watermelon. Both its property of having green skin and its property of having red flesh are factual properties. However, if I can only see the skin from where I stand, then only the property of having green skin can be represented by the factual contents of my experience. As another example, suppose I saccade to Carl's nose after staring at his neck for a while. If his neck becomes barely visible

---

<sup>50</sup> If two contents are subjectively indistinguishable, then they are introspectively indistinguishable. However, I take the notion of subjective indistinguishability to be broader than that of introspective indistinguishability – after all, introspection may not be the only method to determine the contents of an experience. There may be methods a subject can use to differentiate between two introspectively indistinguishable contents. For some candidates of such methods, see Hopp (2016, sec. 2).

after the saccade, his nose is now represented by the factual contents of my experience, but his neck is no longer so represented.<sup>51</sup>

Condition (2) specifies the circumstances under which numerically distinct experiences can be said to have the same factual content. Condition (2)(a) deals with cases of veridical perception. It incorporates Madary's insight and says that changes in spatial location and visual fixation give rise to changes in factual contents. By contrast, condition (2)(b) handles cases of hallucination. The object of a hallucination cannot be said to bear any spatial relation to the hallucinating subject. Hence, if condition (2)(a) were necessary for all factual contents, hallucinations would not have factual contents. That is not my preferred approach. I take it that factual contents, just like Hopp's intuitive contents, are partly individuated by their first-person features. According to Hopp, if one examines different portions of a brown table, "The intuitive contents involved in grasping that portion which is basking in the sunlight over there are identical with those that would, in another context, present a tan object" (Hopp 2008, 220). What it is like to see a brown object in one context can be exactly the same as what it is like to see a tan object in another context. Given that our notion of a factual content is supposed to capture how experiences represent objects from the first-person perspective, we should acknowledge the possibility that an experience

---

<sup>51</sup> Strictly speaking, Carl's nose is not a factual *property*; the property of being Carl's nose is. There is a metaphysical worry here: shouldn't the parts and properties of an individual be distinguished from each other? While Franz's brain is a part of Franz, Franz's rationality is not a part but a property of him. Indeed, the distinction has to be made, but it shouldn't be too difficult to translate talk of parts into talk of properties. After all, Franz has a heart as one of his parts if and only if he instantiates the property of having a heart. There is a related question worth asking about the case of seeing Carl's nose: what exactly is visible to me in this case? Is it a universal, a trope, a quantity of matter, an individual, an event, a fact, a state of affairs, or something else? Unfortunately, this issue has to be pursued somewhere else.

of a brown object sometimes shares the same factual content as an experience of a tan object. By the same token, whenever we attribute factual contents to a veridical experience, we should also attribute factual contents to a subjectively indistinguishable hallucination. What condition (2)(b) does *not* require, however, is that subjectively indistinguishable experiences have the same object.<sup>52</sup>

Before proceeding, two comments are in order. First, recall my terminological distinction between an entity and an object. An entity is whatever exists in some possible world, while an object is defined relationally: it is an entity that an information-carrier carries information about. Given that there many kinds of entities, there are also many kinds of objects; an object need not be an individual. When I speak of experiences of objects in (Factual Contents), therefore, those objects can be properties and events, etc. Second, condition (2)(b) may remind one of how hallucinations are described by disjunctivists. For example, here's a remark by M. G. F. Martin: "some event is an experience of a street scene just in case it couldn't be told apart through introspection from a veridical perception of the street as the street" (2004, 48). According to Martin, this way of individuating experience-kinds carries the least theoretical burden, so such a conception of experience should be the starting point of our theorizing. I wholeheartedly agree with Martin's claim. However, disjunctivists typically go on to say that experiences depend on a certain relation that is both non-representational and psychological (rather than merely physical or causal): Martin calls it "awareness" (2004, 39); Brewer calls it "acquaintance" (2011, 55). I am of the opinion that such a claim gets things backwards. My preferred view is that acquaintance is made possible

---

<sup>52</sup> The issue of the way contents correlate with objects will be further explored in section 4.3.4.

by representation. As Hamid Taieb puts it in describing Brentano's view, "a cognitive act whose object exists bears a specific relation to it in addition to intentionality" (2018, 3). We can take acquaintance to be this "specific relation", but it would be problematic to explain intentionality in terms of it.<sup>53</sup>

Now we can move on to expectational contents. The postulation of such contents is motivated by the Husserlian idea that an experience represents more than what is represented by the factual contents of the experience. As Madary put it, "The phenomenology of vision is best described as an ongoing process of anticipation and fulfillment" (2016, 26). But obviously, it is not the case that anticipations of any kind can be components of visual perception. According to Madary, the anticipations that can be such components are those of the following kind: "If the subject sees a factual property of the object from one perspective, she will anticipate self-generated movement that gives new perspectives on the object will reveal new perspectives on *that* factual property" (2016, 52). Suppose there is a slice of a watermelon lying on my table and I see it as having a red part. According to Madary's thesis, by undergoing this very experience, I will acquire the expectation that I will continue to see the slice's property of having a red part when I walk to the other side of the table; in addition, I will expect to gain new information about the property from the different perspective that I will occupy when I move there. Such anticipations may or may not conform to the actual properties of the object; if one finds out in subsequent experiences that an anticipation is correct, then one has a fulfilled anticipation. Again, one's visual fixation partly determines what anticipations arise in one's experience.

---

<sup>53</sup> I will revisit this issue when I discuss the relational view in section 4.4.



For example, there is a huge difference between what one would expect before and after one saccades to the face of a friend that is about to walk past one. Madary thus holds that in most cases, two experiences differ in expectational contents if they result from different visual fixations (2016, 78).

The expectational contents considered by Madary are limited to those about what one would see after one's bodily movements. It may be that this restriction is imposed simply for argumentative purposes; I don't see any reason that Madary would be against relaxing the restriction. Allowing a greater variety of expectations to count as the expectational contents of perception requires additional arguments, of course. As I will say more in section 5.5.1, the actions of professional athletes suggest that one may perceptually expect more than what results from one's own movements. A professional tennis player acts differently than a novice player even if they are reacting to exactly the same move by the opponent. One could explain this by saying that experts and novices are motivated to carry out different actions by the same visual experience, but I take the better explanation to be that they have different visual experiences. Expert players are more capable of seeing their opponents' bodily movements as initiating certain types of strokes; as a result, they can anticipate and counter the strokes more efficiently. If so, professional athletes undergo experiences whose expectational contents are not about the consequences of their own actions.

Given these considerations, we can define expectational contents as follows:

(Expectational Contents)

Let  $e$  be an experience that  $s$  undergoes at  $t$ . Then  $e$  has an expectational content if and only if  $e$  has some content  $d$  such that:

- (1)  $d$  represents a factual property  $s$  expects to see at the moment succeeding  $t$ ;
- (2) what  $d$  represents partially depends on what the factual contents of  $e$  represent.

Note that condition (2) is indispensable. Suppose I have never heard of durians or jackfruits, but a friend of mine, Anton, just told me that he was about to show me two jackfruits. After he put a fruit on the table, I asked him whether that was a so-called jackfruit. He answered in the positive. Given his answer and what he told me earlier, I now expect to see another jackfruit at the succeeding moment. In this case, my expectation that I will see a second jackfruit is *not* an expectational content of my experience of the first jackfruit. It is a belief. If Anton had shown me a durian, I would still expect to see a jackfruit. The expectation is thus independent of the factual content of my experience, which violates condition (2).

#### 4.3.3 *How Many Kinds of Contents Are There?*

We have now covered the central ideas of HDC.<sup>54</sup> It is time to revisit the two questions raised at the end of section 4.3.1; this will clarify how my formulation of HDC differs from the Husserlian theories I draw on. Let's begin with the first question: what is the relation between the factual and expectational contents of perception? Is one of them in

---

<sup>54</sup> HDC may not be intuitively appealing to some: they might find it hard to accept the claim that expectations can be genuinely perceptual. As I see it, the issue is best discussed in connection with event perception, so I will wait until the next chapter to examine the related arguments. For now, I will simply assume that the claim is true.

some sense derivative of the other? Madary seems to think that factual contents are expectational contents of a kind. Madary does have a notion of factual contents: they are one's representations of factual properties (2016, 44). However, Madary goes on to suggest that the factual content of an experience reduces to its expectational or AF content: "We represent factual properties through visual anticipation of the way those properties will appear within the context of a particular perceptual episode" (2016, 61). If AF contents are the only means by which we represent factual properties, then factual contents reduce to AF contents. To be sure, such a reading may be too strong. Madary explicitly says that he has no business in such reduction: "Visual anticipations depend on there being a representation of a factual property, not the other way around" (2016, 61). But in a later passage, he argues *against* the claim that "there must be something more to visual content than AF content" (2016, 65). It seems that the only way to reconcile Madary's various claims is to adopt the following view. Some representations of factual properties are perceptual contents but the others are not. The latter are what AF contents depend, while the former just are AF contents. The consequence is that the content of an experience is exhausted by its AF content, which "*is the content of the visual anticipations themselves*" (2016, 60). In other words, the content of an experience consists entirely of anticipations about factual properties. Whatever is not an anticipation has no role to play in conscious experiences.

I am not sure that this is Madary's considered view. But if it is, I am worried that it doesn't do justice to phenomenological data. There seem to be many experiences whose contents are not exhausted by expectations, and Madary himself has discussed one such case. Consider the first visual experience one has upon waking up. According to Madary, while some might think that this case is a counterexample to his view, it is not. He argues that one

would have visual expectations of what one would see the moment one opens her eyes. Such expectations constitute the content of one's experience upon waking (2016, 40).

Unfortunately, Madary's response is inadequate. Suppose I open my eyes and feel shocked by the bewildering and rapidly changing scene I see. Since what I see completely goes beyond my expectations, it doesn't make much sense to describe what I see in terms of expectations.

Madary could try to relate the weird experience to expectations. He could describe the experience as the frustration of my expectations. But while fulfilled expectations may be sufficient for representing factual properties, frustrated expectations alone are certainly not. While an expectation can only be fulfilled in limited ways, it can be frustrated in indefinitely many ways. If I slice a watermelon in half and expect to see red flesh inside, my expectation will be frustrated if I see yellow flesh inside, see nothing inside or see a bag of diamonds inside. If so, the very fact that my expectation to see red flesh is frustrated doesn't carry much information about what my environment is like – the fact hardly tells me anything about the factual properties instantiated by the objects in view. Are there other expectations in my experience that might account for the representation of factual properties? I might have the vague expectation that some unfamiliar objects will soon show up in front of me. However, given how rapidly the environment is changing, I don't really have much idea about what to expect. As a result, my expectations of what are about to happen are too impoverished to capture my present environment. Hence, if Madary's view is that one's environment is entirely represented by perceptually based expectations, he would have to say that in the case at hand, my present environment is barely represented by my experience. That is not the right description of my experience. An even trickier issue is that taking

perceptual contents to be exhausted by expectations leads to conceptual difficulties. For we could ask: what is it that fulfills the visual expectations that arose at the moment just past? The answer cannot be present expectations, as our example shows. If Madary is unwilling to accept contents that are not expectations, the only remaining candidate seems to be certain present representations that are not components of conscious experiences. Maybe these are representations employed by the perceptual system but not available to consciousness, or maybe they are automatic associations of a kind more similar to stereotypes. This last resort implies that there are no fulfilling *experiences* – all fulfilling representations are unavailable to consciousness.

The last resort is quite unattractive. The awareness that one's expectations have been fulfilled by one's present experience is an important phenomenon, and it is not a rare phenomenon.<sup>55</sup> The last resort would entail that such awareness is an illusion. I am not ready to accept such a radical view. After all, there is an easy way out of the predicament: we can simply accept the claim that “there must be something more to visual content than AF content” (Madary 2016, 65). The “something more”, I submit, is the factual content of an experience. Therefore, though my formulation of HDC draws heavily on Madary's work, there are two claims that I endorse but Madary probably would not. First, whereas Madary seems to think that factual contents are a species of expectational contents, I take these contents to differ in kind. Second, while I give a disjunctive definition of factual contents, it is unclear that Madary would be willing to do so. If Madary takes factual contents to be

---

<sup>55</sup> For additional discussion of such awareness, see section 4.3.4, where Husserl's notion of *continuous synthesis* is examined.

expectational in nature, then he would have to adopt a disjunctive conception of expectations to accommodate a disjunctive definition of contents. I am not sure how this should be done or whether he would want to do so at all.<sup>56</sup>

#### 4.3.4 *Are Experiences Object-Involving?*

Let's turn to the second question raised in section 4.3.1: on what conditions do two experiences have the same contents? More specifically, can experiences of different objects have the same contents? This issue is a central concern of the Husserlian view held by Smith, which is further developed by Hopp. Their view on this issue is quite straightforward: if two experiences represent different objects, it is impossible for them to have the same contents. Smith gives a Husserlian defense of what he calls *de re externalism*: "the particular object of a perceptual experience is essential to that experience *in virtue of that experience's content*" (2008, 315; see also 314). Similarly, Hopp endorses a thesis that he calls *object determination* (OD): "Necessarily, if two experiences have the same content, then they (re)present the same

---

<sup>56</sup> Hopp also has an argument for counting factual or intuitive contents among the contents of perception. However, his argument doesn't aim to establish that intuitive contents differ in kind from expectational contents; rather, its goal is to show that the factual components of perception should be regarded as sufficient for contents. The argument highlights the epistemic role of perception. According to Hopp, "perceptual and other intuitive acts, including hallucinations, imaginings, dreams, and pictorial memories, have a distinctive sort of phenomenological character that mere thoughts, as such, do not possess" (2010, 7). Hopp calls what distinguishes intuitive states from mere thoughts their *intuitive character*, and the intuitive content of an experience just is the content the experience has in virtue of its intuitive character. Here Hopp takes as his point of departure the Husserlian idea that intuitive states are those capable of verifying thoughts on the basis of their first-person features. The stronger claim Hopp is making is that these first-person features constitute contents of a kind, because "only states with intentional content, and, specifically, a kind of intentional content that represents *how things are*, can be reason-giving" (2010, 10). Since perception is indeed reason-giving, it should be regarded as having contents that represent the way things are.

object” (2011, 173).<sup>57</sup> Let’s say that an experience *has object-involving contents* if and only if it satisfies OD.<sup>58</sup> I am inclined to think that perceptual contents are *not* object-involving, and this is where I depart from the Smith-Hopp account. To clarify what is being contested, we need to discuss a Husserlian idea central to their account.

We have been working with the Husserlian idea that the expectational or horizontal content of an experience can be fulfilled by the experience that follows. As our discussion of Madary’s view suggests, what fulfill expectational contents are not more expectational contents; rather, the expectational content of an experience is fulfilled by the factual or intuitive content of a later experience. Smith notes that for Husserl, such fulfillment enables the subject to undergo a distinctive kind of experience. Husserl calls the experience “the synthesis continually joining the manifold perceptions” or “continuous synthesis” (Husserl [1973] 1997, 132); Smith describes it as “an *experience of identity*” (A. D. Smith 2008, 326).

Suppose I see my wallet at  $t_1$  and want to make sure that there is cash inside, so I open it and

---

<sup>57</sup> There are different ways to read OD. For my purposes, I will assume the following reading. Suppose:

- $E$  is a set of experiences;
- $C$  is a set of contents;
- $O$  is a set of entities;
- For any experience  $x$  and content  $y$ ,  $Bear(x, y)$  holds if and only if  $x$  is the bearer of  $y$ ;
- For any experience  $x$  and object  $y$ ,  $Rep(x, y)$  holds if and only if  $x$  represents  $y$ .

I take OD to be the following conditional:

$$(\forall e \in E)(\forall c \in C)(\forall o \in O)(Bear(e, c) \wedge Rep(e, o) \rightarrow \Box(Bear(e, c) \rightarrow Rep(e, o))).$$

In other words, if  $e$  bears the content  $c$  that represents  $o$  in world  $w$ , then in every world accessible from  $w$ ,  $e$  represents  $o$  insofar as  $e$  bears  $c$ .

<sup>58</sup> The term “object-involving content” has been used in the literature, but I am not entirely satisfied with the way it is typically defined. For example, Susanna Siegel characterizes the object-involving content of an experience as such that “the object seen... is itself part of the content” (2006, 364). I don’t think that objects are ever part of contents.

see dollar bills inside at  $t_2$ . In this case, the expectational content of my experience at  $t_1$  represents the wallet as containing cash, which is then represented again by the factual content of my experience at  $t_2$ . The later factual content thereby fulfills the earlier expectational content. When this happens, one experience seamlessly transitions to another, and my experiences appear to me as jointly belonging to a single coherent course of experience. In this sense, my experiences undergo “synthesis”: they combine with each other into a more complicated representation. Such a representation not only makes the wallet perceptually available to me at both  $t_1$  and  $t_2$ ; it also makes me aware that the wallet I now see at  $t_2$  is the same one as the wallet I saw at  $t_1$ . For this reason, the complex representation is an experience of the identity of the wallet: it represents the wallet as remaining identical across different times.

Since the term “synthesis” may have misleading connotations, I will use the term “experience of continuity” in place of Husserl’s “continuous synthesis”.<sup>59</sup> Hopp cites considerations about experiences of continuity to defend the claim that perceptual experiences have object-involving contents. In his example, one sees a steel ball,  $A$ , which is qualitatively identical to ball  $B$ . What does it take for my visual experience to represent  $A$ , rather than  $B$  or any other qualitatively identical ball? Hopp’s proposed answer utilizes the

---

<sup>59</sup> Some might think that the term “synthesis” carries the connotation that the experiences in question are *deliberately* combined by the subject. That makes the idea of continuous synthesis dubious: it is certainly not the case that one becomes aware of multiple experiences and decides to put them together. This is not what Husserl has in mind. As noted by Smith, Husserl allows some kinds of synthesis to be passive, including those involved in perception. To use the term “synthesis” in this context is just to emphasize that perception requires more than “a mere succession of data” (A. D. Smith 2003, 128). Without contributions from the perceiver, objects cannot appear in perception the way they do.



notion of path connection. Consider arbitrary experiences  $e$  and  $f$  and a possible perceiver that undergoes  $e$  exactly when and where  $e$  occurs. If there is a spatiotemporal path such that the possible perceiver could move along that path and come to undergo  $f$  at the endpoint of that path, then that path is a *connecting path* between  $e$  and  $f$ . Two experiences are *path-connected* if and only if there is a connecting path between them.<sup>60</sup>

According to Hopp, for two experiences to represent the same object, they have to satisfy one of the following conditions, depending on whether they are path-connected (2011, 182). First, if  $e$  and  $f$  are path-connected, then they must be *synthesizable*. That is, if a perceiver moves along a connecting path between  $e$  and  $f$ , thereby undergoing a series of perceptual experiences, then an experience of continuity must co-occur with every experience in the series except the first one. For example, suppose that at  $t_1$ , I expect to see the right wing mirror of my car when I walk past the parking meter occluding it. If I do see it at  $t_2$ , then there is an experience of continuity co-occurring with my experience at  $t_2$ . Further suppose that my experience at  $t_2$  has the expectational content that if I stay where I am, I will continue to see the wing mirror. If this is indeed the case at  $t_3$ , then there is another experience of continuity co-occurring with my experience at  $t_3$ . It follows that my experiences at  $t_1$ ,  $t_2$  and  $t_3$  are synthesizable with each other. Second, if  $e$  and  $f$  are not path-connected, then they must be *harmonious* with each other. This means that the judgments  $e$  and  $f$  are capable of justifying must have the same objects as their subject matter. For example, take my judgments that the London Eye is white and that the River Thames is flowing. Suppose my judgments are justified by two perceptual experiences that successively

---

<sup>60</sup> The terms “path connection”, “connecting path” and “path-connected” are mine.

occur. Despite their temporal proximity, these experiences are not harmonious with each other; the judgments they justify have numerically distinct objects.

Now we can answer what it is that makes my present experience an experience of  $A$  rather than an experience of some other qualitatively identical ball. Take all the possible experiences synthesizable or harmonious with my present experience; call it the *perceptual manifold* of  $A$ . According to Hopp, if an experience represents a ball that is qualitatively but not numerically identical to  $A$ , then the experience cannot possibly belong to the manifold of  $A$  (2011, 182–83). Consider  $B$ , a ball qualitatively identical to  $A$ . Hopp argues by contradiction that no possible experience belongs to both  $A$ 's and  $B$ 's manifolds at once. Assume to the contrary that there is such an experience; call it  $e$ . It follows that for any  $a$  and  $b$  in the manifolds of  $A$  and  $B$ , there are connecting paths between  $e$  and  $a$ , on the one hand, and between  $e$  and  $b$ , on the other. In addition, in both series of experiences, every experience except the first one is accompanied by an experience of continuity.

There are two scenarios in which these are true (Hopp 2011, 184). First, the perceiver simultaneously sees  $A$  and  $B$  and constantly keeps them in view. In that case, while the experiences are indeed synthesizable, they are not internally harmonious. Any experience in the series is capable of justifying two kinds of judgments: those about  $A$  and those about  $B$ . It thus makes more sense to regard each of the experiences as a fusion of two experiences that respectively represent  $A$  and  $B$ . This means that  $e$  is harmonious with experiences of  $A$  if and only if it is *not* harmonious with experiences of  $B$ . We have to conclude that  $e$  simply

doesn't exist.<sup>61</sup> Second, the perceiver keeps *A* in view until she sees *B*, at which point she disregards *A* and attends to *B*. Call this moment *t*. Now, the perceiver's expectation at *t* is presumably that *A* will remain in sight. Since the perceiver shifts her attention to *B* immediately after *t*, her earlier expectation is not fulfilled. As a consequence, no experiences of continuity occur. Since *t* is the only moment at which *e* could possibly occur in this scenario, we have to again conclude that *e* doesn't exist. On the assumption that these two scenarios are exhaustive, any possible experience in the perceptual manifold of *A* fails to belong to the perceptual manifold of *B*.

The observation that perceptual manifolds are mutually exclusive suggests a reasonable criterion of object-involving contents: an experience has object-involving contents about *A* if and only if the experience belongs to the perceptual manifold of *A*. As noted by Hopp, "manifolds exist whether any actual consciousness of the relevant object is carried out; an unexplored cave, for instance, has its perceptual manifold. For this reason, we can understand manifolds as belonging to objects themselves" (2011, 180). My experience represents the cave if and only if it is a member of the cave's manifold. Given the criterion, the inference to OD is straightforward. According to Smith, whether undergoing two successive experiences is sufficient for undergoing an experience of continuity is a matter of the contents the experiences have (2008, 325). The joint sufficiency for experiences of continuity, or the lack thereof, determines whether the experiences are members of the same

---

<sup>61</sup> My reasoning here doesn't quite follow Hopp's own. Hopp appeals to "difference-consciousness" in this part of his argument (Hopp 2011, 184). Though the idea is straightforward, I am not sure how undergoing difference-consciousness violates the conditions of synthesizability and harmoniousness. Insofar as one keeps both *A* and *B* in view, one's earlier and later experiences are synthesizable with each other. This is the case whether or not one undergoes difference-consciousness.

perceptual manifold, which in turn determines their objects. Together with the criterion of object-involving contents, Smith's claim entails OD. It follows that the contents of perception are object-involving; in addition, they would not be so without the joint contributions of both factual and expectational contents.

As noted by Hopp himself, however, the conditions of synthesizability and harmoniousness may *not* be sufficient for object-involving contents. Hopp discusses this worry by presenting a somewhat complicated thought experiment about an eccentric neuroscientist (2011, 186). It seems to me that the worry can be highlighted with simpler cases: those involving a switcheroo.<sup>62</sup> Suppose Anton was showing me a magic trick. After laying several playing cards on the table, he asked me to take a look at one of them without telling him what card it was. I picked a card, identified it as the five of spades and put it back. Then I followed his instruction to stare at the card while he pulled off some seemingly irrelevant tricks. When I was asked to take a look at the card again, I realized that unbeknownst to me, the card had been replaced by the ten of clubs. Now consider the perceptual experience I underwent just before the ten of clubs was revealed; call it  $e_{10C}$ . This experience was synthesizable with  $e_{5S}$ , which was my initial experience of the five of spades. Given that my expectation to see the five of spades was frustrated only after I underwent  $e_{10C}$ , every experience I had between  $e_{5S}$  and  $e_{10C}$  co-occurred with an experience of continuity. Moreover,  $e_{5S}$  was harmonious with  $e_{10C}$ , because for all I knew, I had been tracking one and the same card. If, by undergoing  $e_{5S}$ , I was justified in making the judgment that the card was the five of spades, then I was no less justified in doing so by undergoing  $e_{10C}$ . Despite the fact

---

<sup>62</sup> Hopp himself appeals to such cases in arguing against rival views. See Hopp (2011, 180).

that  $e_{55}$  and  $e_{10C}$  were synthesizable and harmonious with each other, they were about different objects. We are then forced to conclude that the contents of  $e_{55}$  and  $e_{10C}$  were not object-involving.

Hopp does have a solution to this problem. He argues that it is not enough for an experience to be synthesizable and harmonious with the *actual* experiences that precede and succeed the experience. It is also required that the experience “belong to a ‘stable system’ of experiences”, where this means that “in situations not too remote from the actual one, had things been different, my experiences would have been correspondingly different as well” (2011, 186). Imagine that a UFO appeared after a playing card I was looking at disappeared. If my perceptual system had functioned in such a way that I continued to see nothing but a playing card after the UFO appeared, then my perceptual system would have failed to be functional. In that case, I couldn’t be said to have seen the playing card in the first place, whether or not the playing card was later replaced by a UFO. My experience of the playing card thus failed to meet Hopp’s stability requirement. Despite the plausibility of the requirement, however, I am afraid that it doesn’t really address the case of Anton’s show. In situations that were largely similar to the actual one, my experience of the two face-down cards would have remained identical to my actual experience. I would not have become aware that one of them had been replaced by the other. Moreover, if my perceptual system was functioning normally, what happened was precisely what was supposed to happen; Anton’s tricks were designed specifically to trick people with normal vision. For Hopp’s solution to work, an argument must be made that there are reasons to take the apparent synthesizability and harmoniousness in the case under consideration to be deceptive. I am not sure that there are such reasons.

As I see it, even if the stability requirement could somehow resolve the difficulties involved in the case of Anton, there are still reasons to think that perceptual contents are *not* object-involving. Consider again Hopp's example of the qualitatively identical steel balls *A* and *B*. Suppose that there are no possible worlds in which they appear in the same spatial region. It follows that no one ever sees *A* and *B* together, so no one is able to tell *A* from *B*. As a consequence, whether or not one's perception is stable, one's perceptual experiences cannot enable one to respond to *A* and *B* differently. To highlight the issue, we could even allow that in some worlds, some people undergo successive experiences of *A* and *B* that are only separated by a blink. This presumably happens when *A* and *B* are quickly switched with each other; let's assume that they are switched by a machine. Since *A* and *B* are qualitatively identical, even the people in question fail to notice the switch. As a result, experiences of continuity always co-occur with those people's experiences of *A* and *B*. I take this case to be a counterexample to the claim that perceptual contents are object-involving. In this case, one's experiences of *A* and *B* are synthesizable, harmonious and stable; yet nothing in these experiences provides any clue about which of *A* and *B* one is seeing. It thus makes more sense to say that in this case, one's experiences of *A* and *B* do not have object-involving contents.

There is a radical response to my argument: one could argue that the experiences in question fail to have contents. If so, OD is still true for any experience that does have contents. I don't think this response works. There is nothing unusual about the experiences under consideration; their similarity to everyday experiences justifies the claim that they have contents. Another response is that the scenario just described is metaphysically impossible. In other words, the sentence "necessarily, steel balls *A* and *B* do not show up in the same

spatiotemporal region” is a contradiction. But is it? If we translate the sentence into first-order modal logic, it is quite easy to come up with a model that satisfies the sentence. Define a two-place predicate  $CLxy$  such that  $CLxy$  is satisfied at world  $w$  if and only if there is a spatiotemporal region in  $w$  such that  $x$  and  $y$  are co-located at the region. We simply define the extension of  $CL$  to be such that in every world, the ordered pairs  $\langle A, B \rangle$  and  $\langle B, A \rangle$  are not in the extension of  $CL$ . This would yield a model that satisfies the sentence, which means that the sentence is not a contradiction. One could still argue that this model doesn’t correspond to anything in reality, and the debate would then become one about the criteria of metaphysical possibility. I don’t have anything useful to say about that issue. I can only point out that if my observations are correct, OD is a very strong thesis about metaphysics: it imposes substantial constraints on what kinds of worlds are possible. While there might be considerations in favor of such constraints, I am not ready to accept OD before those considerations become clear.

#### 4.3.5 *HDC: A Second Formulation*

I have argued that perceptual experiences do not feature object-involving contents. My argument is essentially that some perceptual experiences of individuals may be counterexamples to OD. Does this mean that perceptual experiences never represent any individuals? I don’t think such a conclusion is warranted. Perceptual experiences do represent individuals; it’s just that even if two experiences represent different individuals, their contents might still be the same. The best way to flesh out this view, I submit, is to

consult Campbell's recent analysis of perceptual selection.<sup>63</sup> As we will see, doing so supplies the missing piece to my formulation of HDC.

In a recent book co-authored with Quassim Cassam, Campbell argues that the perceptual relation holds “between a thinker and an array of visible properties at various locations, available for use in the selection of objects as figure from ground” (2014, 64–65). This view can be illustrated with an example. Suppose Anton is presented with a photo taken in Egypt, which shows the Sphinx with several pyramids on the horizon. He is asked to determine whether there are pyramids in the image. There will be a very brief phase in Anton's experience in which well-individuated objects have not yet appeared. Over the first few hundred milliseconds, Anton might be aware of nothing but various environmental features located at different locations: some brown color patches, triangles and irregular shapes here and there, etc. After that, given that Anton's task is to find pyramids, Anton might be able to visually attend to the triangles in the image. The important point here is that to attend to *triangles* is not yet to attend to *triangular individuals* – in this case, the pyramids. To attend to triangles is to become aware that the property of triangularity is instantiated in the environment, but to attend to triangular individuals is to perform figure-ground segregation and to single out the objects that are the bearers of triangularity. Though attending to the

---

<sup>63</sup> This may seem a strange choice. Campbell is one of the most prominent proponents of the relational view, but from the very beginning, I have stressed that HDC is a representational view. Despite Campbell's stance, however, I don't think his analysis of visual selection presupposes the relational view. I will therefore integrate his analysis into a representationalist framework. Campbell would not like this; he argues that visual selection is best understood in a relationalist framework. I don't think his arguments succeed, so I will raise objections to his arguments in section 4.4. For now, I will set this issue aside.



triangles in the image enables Anton to attend to the triangular individuals there, the former could occur without the latter.

So far we have only considered Anton's awareness of environmental features. How does the awareness of individuals emerge? To answer the question, Campbell draws the distinction between *selection* and *access*: "Grabbing the thing out from its background (selection) is one thing, and characterizing it (access) is another" (2014, 54). Once Anton becomes aware that the environment depicted by the image contains triangles, he could then fixate on one of the triangles and examine the individual bounded by it. In doing so, he *selects* the individual bounded by the triangle on the basis of the triangle. This nevertheless doesn't tell us which properties of the individual Anton will *access*. When examining the individual, he could become attracted to the texture of the individual and direct his attention accordingly; in that case, only the texture of the individual is accessed. But it is also possible for him to access the color alone or to access both the color and the texture. Most importantly, Anton need not even access the triangular shape of the individual despite the fact that the individual is selected on the basis of its triangular shape. After all, after singling out the individual to be examined, Anton could focus on any feature he is interested in; if he is asked to determine, say, whether any text is superposed on the triangles in the image, there is no need for him to examine the shapes of the targets once the targets have been selected.

Campbell argues that there is a close connection between the experience of an individual and the experience of the property in virtue of which one can visually select the individual. According to Campbell, in cases where "experience of object O is causally made possible by experience of property F", it makes sense to infer that "property F will constitute the visual 'mode of presentation' of the object O" (2014, 53; also see 67). Suppose that one

second after the image is shown, Anton is able to complete the task and identify a pyramid in the image. Campbell's analysis would be that Anton sees the pyramid under the mode of presentation of a triangular object. The reason is that Anton visually selects the pyramid on the basis of triangularity; the experience of the latter thus causally makes possible the experience of the former. Note that the relation of "causally making possible" is supposed to disambiguate certain experiences similar to each other. Consider two cases in which I see a watermelon. In the first case, I select it on the basis of its color and then access its shape; here what I see can be described as "the green object is round". In the second case, I select the watermelon on the basis of its shape and then access its color; what I see in this case can be described "the round object is green". Can we say that I undergo the same experience in these two scenarios? Campbell would give a negative answer (2014, 64). In the first case, my experience of the watermelon is causally made possible by my experience of its green color. My visual selection of the watermelon is caused by my awareness of its green color, and such a causal fact distinguishes my experience in the first case from that in the second case.

Couched in representational terms, Campbell's analysis of visual selection suggests a way to make sense of my earlier claim: perceptual contents that are not object-involving may nonetheless represent individuals. On the representationalist reading of Campbell's analysis, when we look for linguistic expressions that describe perceptual representations of individuals, the best candidates are definite descriptions. We may describe Anton's experience as representing the pyramid under the mode of presentation "the triangular object located at location  $x$ ", where  $x$  could be specified either in an egocentric coordinate system, i.e. relative to one's own body, or in an allocentric one, i.e. relative to external

objects.<sup>64</sup> Obviously, different objects could be represented under the same mode of presentation. Suppose there are multiple pyramids in the image Anton is looking at, and two of the pyramids, *A* and *B*, are at roughly the same distance from the Sphinx. Anton visually fixates on *A* and perceives it as a triangular object a few hundred yards away from the Sphinx. Franz is also looking at the image, but he focuses on *B* and sees it as a triangular object a few hundred yards away from the Sphinx. Here Anton's and Franz's experiences feature the same mode of presentation, but their experiences represent different individuals. If so, though their experiences lack object-involving contents, the experiences still represent individuals.

Note that on this representationalist reading of Campbell's analysis, it is *not* the case that an experience has some indexical component responsible for picking out its object in a given context. Definite descriptions are not indexicals; saying "the pub on Beacon Street" is quite different from saying "that". As a result, the present view is immune from the various objections Hopp raises against explaining the perceptual representations of individuals in terms of indexicals. Hopp is worried that an indexical account not only gets things backwards but also fails to provide any real explanation. On the one hand, the referent of an indexical is fixed by one's perception of an object, not the other way around (2011, 124–25). On the other hand, an indexical like "that" can be appropriately used no matter what one is seeing, but no content of perception should have that kind of flexibility (2011, 120–23). Explaining the latter in terms of the former thus amounts to a failure of explanation. The present view is susceptible to neither objection. If an experience represents an object under

---

<sup>64</sup> For a discussion of the two kinds of coordinate system, see Gramann et al. (2006).

the mode of presentation “the triangular object a few hundred yards away from the Sphinx”, this very fact is capable of explaining how one can utter “that is a pyramid!” and achieve referential success with the indexical “that”. Furthermore, while one can appropriately refer to an object with “that” regardless of the experience one is undergoing, one cannot do so with a definite description. If one only sees a blackboard, it makes no sense for one to think “the triangular thing over there is huge”.

Two caveats are in order. First, while the individual-representing contents of perception may be described by using definite descriptions, the relevant definite descriptions are not *Russellian* ones.<sup>65</sup> In Russell’s theory, a definite description fixes its referent by specifying certain conditions that are uniquely satisfied by something in the domain of quantification. That is not the means by which Anton’s experience is directed at its object. Though Anton’s experience represents a pyramid under the mode of presentation of a triangle, the mode of presentation doesn’t function to relate Anton to any uniquely instantiated property in the environment; there are many triangles in the image. Indeed, if the relevant modes of presentation were Russellian definite descriptions, it would be hard to see how Anton and Franz could represent different pyramids in the image under the same mode of presentation. Second, it is worth emphasizing that the definite descriptions we have been using are linguistic descriptions of perceptual contents *and nothing more than that*. The very contents of perception are not structured linguistically, so they cannot literally be definite descriptions. They could be Fregean senses, as suggested by, for example, Brad

---

<sup>65</sup> For some difficulties facing an account of perception inspired by Russell’s theory of definite descriptions, see Pylyshyn (2007, 15 fn. 4).

Thompson (2009), but only if we understand Fregean senses in a somewhat non-Fregean way: perceptual senses could still be modes of presentation of a kind, but their compositional semantics will be quite different from that of linguistic senses. In particular, perceptual senses should not be allowed to compose propositions, as discussed in section 4.2.1.

We have emphasized that the representation of a triangular individual cannot be identified with the simpler representation of the triangular shape of the individual. But how exactly should we describe the phenomenological difference between the two experiences? Other than that there are causal relations between the two, Campbell doesn't tell us much. This may make one suspicious of the distinction between the representation of a triangle and that of a triangular individual; there is surely a verbal distinction, but does the distinction make any sense from the first-person perspective? My proposal is that while the former is a representation that aims at a trope, the latter is a representation that aims at a *foundational system*, to put it in a term of Peter Simons'. A foundational system is a bundle of tropes that "can in principle exist alone except for its dependence on its constituent trope parts, and on any larger wholes of which it is a dependent part" (P. Simons 2000b, 148).<sup>66</sup> Take a black key of a piano; call it *k*. Since *k* cannot exist without having a color and a shape, it depends on the color and shape tropes that are constitutive of it.<sup>67</sup> There is a sense in which *k* also depends on the piano: it could not be what it is, i.e. a key of the piano, if the piano didn't exist. But *k*'s dependence on the piano isn't quite the same as its dependence on the tropes.

---

<sup>66</sup> Actually, the definition of a foundational system is more complicated than that; it is defined in terms of a relation Simons calls *foundational relatedness* (1994, 562). The definition cited here is a simplified one, which, for my purposes, is sufficient.

<sup>67</sup> In Simons' account, such dependence is specific, not individual. For a discussion of the distinction, see section 3.5.2 of this dissertation.

If  $k$  is detached from the piano, it continues to exist, but there is no way to detach the color and shape tropes from  $k$  without destroying it. Therefore, while  $k$  is a dependent part of the piano, it could in principle exist alone insofar as its constituent trope parts exist. It follows that the key of the piano is a foundational system. Needless to say, the piano itself is another foundational system.

On my proposal, a perceptual experience picks out an individual by picking out a foundational system to which the basis of selection belongs. For example, when Anton sees the triangular shape of the pyramid, he sees a triangle trope. On the basis of this visual experience, he is able to visually select a foundational system of which the triangle trope is a member – here the foundational system is of course the pyramid. There is a phenomenological difference between the experience of a triangle and that of a triangular individual because what is experienced in the former is only a member of what is experienced in the latter. Now, to represent a foundational system, Anton's experience has to represent some members of the system other than the triangle trope. His experience has to represent some tropes with which the triangle trope must co-exist. How are those tropes represented? I suggest that they are represented by the expectational contents of Anton's experience. Upon seeing a triangle, Anton comes to expect the existence of other environmental features. For example, he expects that as the experience unfolds, he will be able to tell whether the triangle is actually a three-dimensional tetrahedron rather than a two-

dimensional figure.<sup>68</sup> To describe Anton's expectation, we could say that he expects the triangle to be an entity that has a volume; he thus expects a volume trope to co-exist with the triangle trope.<sup>69</sup> Once such an expectation is fulfilled by the factual contents of his experience at the succeeding moment, he acquires new expectations. He expects that, with more time passing, the surface texture of the tetrahedron will be more clearly seen. In other words, he expects a texture trope to co-exist with the triangle and volume tropes. Such expectations differentiate Anton's experience of the pyramid from his experience of the triangle. When one undergoes experiences in a less familiar environment, the expectational contents of one's experience could be more impoverished. Even in those cases, however, one represents an individual in addition to an environmental feature only if one represents certain tropes that bear dependence relations to the environmental feature.

The present proposal entails that, without perceptually based expectations, Anton's initial experience of the pyramid would be exactly the same as his experience of the triangle. The experience of an individual occurs when one imposes one's expectations on a perceived trope.<sup>70</sup> Does it follow that the individuals represented in perception are nothing but our projections? By no means. There is indeed a sense in which one's experience is subjective

---

<sup>68</sup> I do not mean to suggest that Anton must possess the concept of a tetrahedron. While the opponents of nonconceptual contents might take such a concept to be necessary for any experience of a tetrahedron, I am of the opinion that one can tell a tetrahedron apart from a two-dimensional figure even if one has had no exposure whatsoever to the idea of a tetrahedron.

<sup>69</sup> This doesn't mean that Anton himself would describe his experience this way. Very few people would ever use the words "entity" and "trope" to describe what they are seeing. Even so, when we theorize about Anton's experience, there is nothing objectionable about such a description. In everyday life, few would describe what they see as "sense-data" or "surface reflectance properties", but this doesn't prevent philosophers from using these terms to describe what is experienced in ordinary life.

<sup>70</sup> Of course, "impose" is just a figure of speech. What expectational contents an experience has is in general beyond the perceiver's control.

when one perceives an object with the help of perceptually based expectations. But such expectations can be entirely veridical. If Anton expects to see stone-like texture when he examines the surface of the individual bounded by the triangle, his expectations carry accurate information about the pyramid. In fact, on the present proposal, we can specify the veridicality conditions of an individual-representing experience in a straightforward manner. Upon visually selecting a pyramid on the basis of a triangle, Anton undergoes an experience of the pyramid that is caused by an experience of the triangle. I suggest the following veridicality conditions of the former: it is veridical if and only if Anton expects to see a foundational system that is a pyramid, on the one hand, and the triangle used to select the foundational system is indeed a member of a pyramid, on the other. How do we determine whether the foundational system Anton expects to see is a pyramid? We do so by considering the tropes Anton expects to see upon seeing the triangle trope. If those tropes are indeed capable of jointly composing a pyramid with the triangle trope, then the foundational system Anton expects to see is a pyramid. Otherwise, it is not. However, even if Anton's expectations are indeed about the components of a pyramid, the triangle used by Anton to visually select the pyramid could turn out to be a member of another foundational system, say, a triangular flag that looks somewhat like a pyramid from afar. In that case, Anton's experience of the pyramid fails to be veridical.

To be sure, the claim that perceptual experiences represent foundational systems is far from self-evident. Why think that, in perception, we ever expect any trope to co-exist with a perceived trope? While I don't have any knockdown argument, I think my claim is supported by a variety of everyday experiences. For example, if Albrecht suddenly notices something moving towards him, he will try to dodge even before he can tell what it is. What



contents should we attribute to Albrecht's visual experience to make sense of his action? It won't do to say that his experience represents nothing but an instance of motion. What is moving could be a plastic bag blowing in the wind, which wouldn't hurt Albrecht even if he is hit. Motion itself doesn't pose any threats. The better explanation is to say that Albrecht's experience represents a foundational system that, in addition to a motion trope, may very well contain a solidity trope. The motion trope is perceived to depend on the solidity trope – the motion is perceived as the motion of something solid. For this reason, the perception of the motion trope prompts Albrecht's attempt to dodge. Or consider a wooden book box that looks exactly like a real book. Edmund sees it and mistakes it for a real copy of a science fiction that he wants to read, so he reaches for it. While the factual contents of Edmund's experience may represent nothing more than the cover-looking surface of the book box, he expects there to be text inside. His experience thus represents both a cover trope and a text trope, which are tropes that depend on each other to constitute a book (together with several other tropes). In this case, Edmund's experience represents a foundational system whose members include the two tropes, among others. It then emerges that such examples are not hard to come by. The abundance of such examples in everyday life suggests that the idea of perceptually representing a foundational system is a reasonable one.

I have said that the primary concern of this chapter is the content-object relation in perception. Now that I have given my argument that foundational systems are among the objects of perception, it is time to make my official formulation of HDC explicit:

(The Husserlian Dual-Component View)

A perceptual experience represents an individual if and only if it represents a foundational system by means of its factual and expectational contents.

There are several potential objections to this view. The most serious ones will be those facing all versions of the representational view, i.e. the objections from the relational view. But even if one remains within the representationalist framework, there are at least two phenomenological reasons to criticize my proposal, which incorporates a representationalist reading of Campbell's analysis. First, one could object to the idea that at least some visual experiences feature successive experiential phases such as the mere awareness of environmental features, the selection of an object and the access of the object's properties. One could argue that we always see well-individuated objects whose rich meanings are immediately available to us. There are at least two ways to respond to this objection. First, it is worth noting that the successive phases are supposed to take place over just a few hundred milliseconds. While it is not easy to become aware of these phases in introspection, an abundance of studies in empirical psychology testifies to the fact that the perceptual individuation of objects takes time (see, for example, the papers in Ögmen and Breitmeyer 2006). Or consider the Gurwitschean theory of consciousness proposed by Jeff Yoshimi and David Vinson (2015). Commenting on the famous gorilla experiment conducted by Daniel J. Simons and Christopher F. Chabris (1999), they argue that when one is preoccupied with a given task and fails to notice a gorilla in front of one, one's consciousness of the gorilla may be "a fragmentary qualitative structure which could allow for some sense of a patchy black surface, even when that experience has not been bound into a focal awareness of a gorilla"

(Yoshimi and Vinson 2015, 114). While such experiences may still be representational, they are not representations of individuals as subjects of properties. Second, I think Madary's point about the periphery of the visual field definitively shows that the objection is misguided. What's so powerful about the indeterminacy phenomena is that they can be easily observed in everyday life. If a familiar face appears on the periphery of one's visual field, one simply cannot tell whose face it is if one doesn't saccade to the face. There is no reason to think that such faces are individuated the moment they appear. As such indeterminacy phenomena again confirm, it is often the case that meaningful objects emerge only at the end of an initial experiential process, even if the process is very brief.

One might feel the need to resist the present proposal for another phenomenological reason. Husserl puts much emphasis on what he calls "identity-consciousness": given two perceptual experiences, a state of identity-consciousness "unites the two perceptions and thereby brings about the consciousness of their object as one and the same" (Husserl [1973] 1997, 23). One enters a state of identity-consciousness by undergoing an experience of continuity. As we have seen, this is why Smith describes experiences of continuity as those of *identity*. One could object to my proposal by appealing to this Husserlian idea: one could argue that there is no need to invoke the representations of unseen tropes in a foundational system. What distinguishes Anton's initial experience of the pyramid from his experience of the triangle trope is an experience of identity. Suppose Anton sees the triangle trope at  $t_1$  and visually selects the pyramid on the basis of the triangle at  $t_2$ . The latter happens because Anton becomes aware that the object experienced at  $t_2$  is precisely the same as the one experienced at  $t_1$ ; in this experience of identity, Anton comes to represent a triangular individual that is more than a triangle trope. I have the suspicion that such an alternative

explanation goes in a circle. If Anton undergoes an experience of identity at  $t_2$ , it must be the case that his experience at  $t_1$  had expectational contents about a triangular individual. These expectational contents are fulfilled at  $t_2$ , so Anton experiences the triangular individual as remaining identical through time. But isn't the question precisely how Anton is able to represent a triangular individual? If we assume that Anton's experience at  $t_1$  already represented a triangular individual by its expectational contents, we have failed to provide an explanation. To put it differently, such a "solution" presupposes that Anton's experience at  $t_1$  already represented more than a triangle trope, but that is precisely what is being contested by Campbell's analysis.

Even if we set aside the circularity of the alternative proposal, it is unclear to me that experiences of continuity are really properly described as those of identity. I would think that experiences of continuity come in many varieties and that experiences of identity are just one of them. But as mentioned above, Smith seems to identify the two. Here's Smith's account:

As I walk round an object while keeping my eye on it, formerly hidden aspects come into view. Their sensory presence 'fills' or 'covers' the earlier empty intentions that were, as implicit elements in the perception's intentionality, empty directed to those parts. The present fulfilled phase of perception is synthesised with the earlier, partially empty phases in a continuing, unbroken sense of the persisting identity of the object. (2008, 326)

Smith's words seem to suggest that we experience identity whenever certain expectational contents are fulfilled by later factual contents. But an experience of continuity just is an experience that occurs when certain expectational contents are fulfilled by later factual contents. If so, it follows that any experience of continuity represents objects as remaining identical through time. That cannot be right. Experiences of continuity occur even when we perceive events, but we often perceive events without tracking any particular individual.

Suppose I see a bowling ball hitting the pins. At  $t_1$ , I expect the pins to fall, and they indeed fall at  $t_2$ . The expectational contents of my experience at  $t_1$  are fulfilled by the factual contents of my experience at  $t_2$ , so an experience of continuity co-occurs with my experience at  $t_2$ . But from the first-person perspective, it doesn't appear to me that I am tracking any particular object and experience its identity across times. I am not paying attention to any particular pin; I might not even focus on the bowling ball. Or suppose I am watching snow fall. I see countless snowflakes, and my expectation that there will be even more keeps getting fulfilled. Despite this, I do not experience the transtemporal identity of any particular snowflake. This is not to deny that I could try to track the trajectory of a snowflake; the point is that the expectational contents of my experience can be fulfilled even if I don't make any such attempts. If so, there are experiences of continuity that do not represent the transtemporal identity of any individual.

The moral, I submit, is that experiences of continuity constitute a genus that has many species. While the perception of individuals and the perception of events may both involve experiences of continuity, they do so in different ways. On my proposal, one undergoes an experience with expectational contents the moment one sees an entity, whether the entity is an individual or an event. The role played by the expectational contents of one's experience nevertheless depends on the nature of the experience. In a perceptual experience of an individual, expectational contents function to represent tropes that co-exist with the tropes presently represented by factual contents. This allows one to represent a foundational system, some of whose members already exist but are not yet in view. By contrast, in a perceptual experience of an event, expectational contents function to represent a future phase of the event that might succeed the phase presently represented by factual

contents. One thereby represents an ongoing event, some of whose temporal parts have not yet come into existence. Hence, when Anton saccades to a pyramid, which is an individual, his expectation may be that he will soon see a texture trope that is already there but not yet clearly seen. In contrast, when Anton sees a school of flying fish jumping out of the water, which is an event, his expectation may be that he will soon see several splash tropes, which constitute a phase of the presently perceived event that has not yet occurred but is about to. It then emerges that the expectational contents of one's experience enable one to perceive entities of different kinds by assuming different roles in different contexts. Since there are several such roles, there are also several forms experiences of continuity may take. If so, not every experience of continuity is an experiences of identity.

We have now seen how my proposal could help to make sense of individual-representing perception. In the next chapter, I will apply the proposal to the case of event-representing perception. But before that, let's consider how HDC might fare in responding to the main rival of the representational view: the relational view.

#### **4.4 Challenges from the Relational View**

One of the most heatedly debated topics in contemporary philosophy of mind is whether perception should be understood in representational or relational terms. As I see it, the best way to frame the debate is found in the work of Matthew Soteriou. According to him, the debate focuses on those conscious perceptual experiences that are veridical. What the relational view denies is that the nature of such an experience "is simply determined by the obtaining of a mental state which has an intentional content with veridicality conditions" (Soteriou 2010, 225). In contrast, what the relational view affirms is that "there are

phenomenally conscious states whose obtaining requires the obtaining of a relation of ‘awareness of’, but which *cannot* be specified independently of that relation” (2010, 234). Suppose I am undergoing a veridical experience in which I see a bowl of ramen. On the representational view, I am aware of the bowl of ramen when undergoing the experience because the experience has a content that represents the bowl of ramen. The relational view takes the opposite route: if my experience has any content at all, it does so because the awareness relation holds between me and the bowl of ramen. The debate between the two views is therefore one about which of representation and awareness is more basic. Obviously, for the debate to be a substantive one, the proponents of the representational view must come up with a way to define representation independently of awareness; similarly, the proponents of the relational view must show us how to understand awareness in non-representational terms.

There is no way to address the large literature that results from the debate in a single section, so my goal here is much more modest than that: I only intend to defend the representational view against two objections raised by Campbell. I focus on these two objections because of their potential relevance to HDC. There are myriad other objections that can be found in the literature on the debate, but they have to be addressed elsewhere as they do not have this much direct bearing on HDC.

Campbell’s first objection aims to show that the representational view gets things backwards. The objection consists of three main claims. The first claim is that in our visual experiences, selection is more fundamental than access (2014, 51). The second claim correlates representation with access: “what the subject is visually representing has to do with the level of visual access, not the level of selection” (2014, 69). The third claim

correlates selection with the awareness relation: “we should be thinking of the visible characteristics of the scene as being there in experience, in the sense that they are available for use as the basis on which we select objects or regions whose characteristics we can go on to access visually” (2014, 72). Here I interpret Campbell as referring to the awareness relation when he speaks of environmental features as “being there in experience”. Together, the three claims entail that the awareness relation is more fundamental than perceptual representation. If so, it is wrong to account for the former in terms of the latter.

The objection fails to refute the representational view because the second claim is rather unmotivated. I think there is a coherent way to think of visual selection as representational as well, which has already been laid out in my presentation of HDC. When I see an environmental feature, my experience represents a trope. Visual selection happens when my experience comes to represent a foundational system that contains the trope. But the representation of a foundational system can be understood in purely representational terms: it is the joint achievement of the factual and expectational contents of my experience. Since nothing in Campbell’s argument rules out this possibility, his conclusion is too hasty.

The second objection from Campbell is directed to a version of the representational view defended by Martin Davies (1997), which Campbell summarizes as saying that “consulting perception will only provide you with a number of existential propositions” (Campbell 2002, 124). Take two qualitatively identical laser pointers, PT1 and PT2. According to the view, the same content is exemplified by both my experience of PT1 and my experience of PT2 – the content that can be expressed by the existentially quantified sentence “there is a black laser pointer in front of me”. Campbell argues that this implies a problematic idea, i.e. the idea that “perception could not provide you with knowledge of



which particular thing you are talking about” (2002, 124). Suppose I look at PT1 and say, “that laser pointer looks like a pen”. I presumably know that the laser pointer I am talking about is the one I am looking at, which just is PT1. However, such knowledge would become inexplicable if the content of my experience can be fully expressed by the sentence “there is a black laser pointer in front of me”. The sentence is true whether the laser pointer in front of me is PT1 or PT2, so it is mysterious how undergoing a mental state whose content is expressed by the sentence enables me to know the referent of “that laser pointer”.

Campbell’s argument is effective against the kind of representational view that conceptualizes perceptual contents as existentially quantified propositions, but a proponent of the representational view need not accept such a conception of contents. As Hopp points out, even if two qualitatively identical objects are presented to me in turn and I cannot tell which object I am seeing now, it doesn’t follow that I do not see the object in front of me (Hopp 2011, 179–80). If so, there is no reason to think of perceptual contents as existentially quantified propositions in the first place – they are more akin to propositions that contain only individual constants. To be sure, the idea that perceptual contents are existentially quantified propositions is intuitively appealing. Its intuitive appeal is nonetheless derived from the problematic assumption that seeing shareable properties and seeing individuals are somehow mutually exclusive. The assumption can be stated more precisely: if (a) one sees an individual and (b) the properties represented by one’s experience is exhausted by the properties the individual shares with some other individual, then one’s experience fails to represent the individual. But this assumption cannot be right. If PT1 and PT2 are of the same color, then one of the things a PT1-representing experience is supposed to do is to

represent the color that PT1 shares with PT2. Representing the color of PT1 enables one to see PT1 instead of preventing one from doing so.

As I see it, the right thing to say about the perceptual experience one undergoes when seeing PT1 is that it represents PT1 but fails to supply one with optimal access to PT1 – one’s perceptual experience cannot represent many of the properties that set PT1 apart from PT2, such as its historical property of being made earlier than PT2. The access to this historical property may be gained if one’s perception is augmented by the instruments in a chemical lab, but one’s bare eyes fail to afford one such access. Regardless, it is simply not true that one’s perceptual experience doesn’t represent PT1 – it is just that one’s experience doesn’t represent *everything* about PT1 that can be represented. Our observations thus suggest that the representation of an individual is not an all-or-nothing matter but a matter of degree. The more properties of an individual one has access to, the better one represents the individual.

This point is well accommodated by HDC. Even if I see nothing but the color trope that partly makes up PT1, I can still visually represent a foundational system whose members include the color trope, thereby representing PT1 itself. However, my perceptual access to PT1 will be improved if I continue to look at PT1 and come to represent more members of the foundational system. Whether or not this happens, I enjoy access to PT1 itself insofar as one of the tropes that PT1 consists of is represented by my experience. If I say, during this process, “that laser pointer looks like a pen”, there is no mystery of how I come to know the referent of “that laser pointer” – my knowledge is acquired via the access to PT1 that is supplied by my experience. HDC is thus immune to Campbell’s second objection.

## 4.5 Conclusion

As stated at the beginning of this chapter, my goal in this chapter has been to defend a version of the representational view that satisfactorily accounts for both the contents and objects of conscious perception. I have argued that that the Husserlian dual-component view does exactly this. The Husserlian notions employed by HDC, i.e. those of factual and expectational contents, are not only phenomenologically tenable but also theoretically flexible. By combining these notions with Simons' idea of a foundational system, HDC suggests a way to make sense of our perceptual experiences of individuals. It is therefore a theory of perception that is worth considering alongside the more mainstream alternatives, such as perceptual propositionalism, perceptual pictorialism and the relational view. In an important sense, however, this chapter is only the first part of my two-part defense of HDC. It is my belief that one of the most valuable features of HDC is its potential to illuminate the nature of event perception – it is to this topic that I will now turn.

## CHAPTER 5

### PERCEPTION: THE EXPERIENCE OF EVENTS

The concluding chapter of my dissertation expands on the Husserlian dual-component view proposed earlier and develops a theory of event perception. I argue that our perceptual experiences of events are structurally different from those of individuals and time in that the experiences of events require a specific kind of awareness: the awareness of temporal boundaries of the kind that individuates an event in accordance with its normal course. My analysis of such awareness makes heavy use of the idea of expectational contents. If the analysis turns out to be sensible, therefore, it gives us additional reasons to think that the Husserlian dual-component view is worth serious consideration.

It may seem unclear why we need a theory of event perception. Don't we all have a decent grasp of such perceptual experiences? A perceptual experience of an event just is one caused by an event; the nature of such an experience is completely determined by the event that causes the experience. What else needs to be said? It is not hard to show that this way of understanding event perception is overly simplistic. A moment's reflection indicates that the nature of an event-representing experience may not be completely determined by its cause. Suppose Mortiz just put his right hand into his pocket. Both Herbert and Wilhelm witnessed Mortiz's action. Herbert is a police officer, and Mortiz's appearance struck Herbert as quite similar to a criminal wanted for firearms trafficking. Wilhelm is an ordinary civilian and didn't notice anything special about Mortiz. Hence, though Mortiz's action could appear to Herbert as the action of reaching for a gun, the action certainly didn't appear that way to

Wilhelm. In this case, Herbert's and Wilhelm's visual experiences were caused by the same event, but it is questionable to say that they had the same visual experience.

To be sure, there are alternative ways to analyze the case just mentioned. One could argue that the difference between what Herbert and Wilhelm experienced was not a *perceptual* one; it is a difference between the judgments made, the beliefs entertained or the concepts applied. To highlight the disagreement between these views and my own, consider an experiment conducted by the psychologist Darren Newtson (1973).<sup>71</sup> The subjects in the experiment were shown a video clip and asked to individuate the events in it. According to the instructions, the subjects had to press a button whenever they perceived the ending of an event and the beginning of another. As it turns out, the subjects broke the recorded process into more events if the recorded process unfolded in a more unexpected way. The experiment thus suggests that there is no straightforward correlation between what events cause one's experience and how one individuates events in one's experience. Newtson's work enables us to devise a thought experiment that brings into sharp focus the phenomenological disagreement between two competing ways to think about event perception:

(The Central Thought Experiment)

Imagine that Alexius and Kasimir are watching the same event unfold in a Newtson-style experiment. Since Kasimir is less familiar with what they are watching, the turns of events

---

<sup>71</sup> The experiment will be examined in greater detail in section 5.3.

are more surprising to him. As a consequence, Kasimir presses the button more frequently than Alexius does. Now we ask: is Kasimir's *perceptual experience* different from Alexius'?

As can be expected, my answer is that their perceptual experiences are indeed different. My answer entails that the nature of an event-representing experience is not simply determined by the event causing the experience – a more complicated account must be given to appropriately describe the way event-representing experiences appear within the first-person perspective. By contrast, my opponents will give a negative answer to the question posed in the Central Thought Experiment. Their answer entails that it is wrong to characterize the phenomenon highlighted by Newton's experiment as a perceptual one. Whatever differences there are between the ways Kasimir and Alexius represent the events shown to them, the differences are cognitive.

The goal of this chapter is to defend my response to the Central Thought Experiment. I will supply the theoretical details needed to substantiate my answer to the question posed in the thought experiment and defend my theory against potential objections. I will argue that Kasimir's perceptual experience differs from Alexius' because the expectational contents of Alexius' experience better represent the structure of the recorded events than those of Kasimir's. As a consequence, whereas Kasimir often presses the button prematurely, Alexius is less inclined to do so. Before I flesh these out, however, I need to explain how my concerns are related to broader considerations. I believe that the kind of awareness highlighted by the Central Thought Experiment is a critical feature of event perception – it is the feature that makes event perception phenomenologically distinctive. But I have not shown this; to do so, I need to contrast our experiences of events with two

related forms of experience, i.e. the experiences of individuals and time. And I have not told the reader why the issue of event perception merits the attention of philosophers. Before I delve into the various forms of perception, therefore, I will say a few words about the significance of event perception for the philosophical study of perception in general.

This chapter is divided into six sections. Section 5.1 makes a case that the neglect of event perception in contemporary philosophy of mind is unjustified. Sections 5.2 and 5.3 examine our experiences of individuals and time. I will argue that our experiences of events have features that are not shared by these experiences, so the former are not reducible to the latter. Section 5.4 presents my own view on event perception, while Section 5.5 makes explicit my response to the Central Thought Experiment. A conclusion then follows, in which I make a brief suggestion on further research.

## **5.1 Why Care about Event Perception?**

Event perception has never been the topic of any major debate in philosophy. I believe this is a mistake. After all, we see events all the time in our everyday life: sports competitions, concerts, summer storms and ocean waves are anything but rare. This fact is acknowledged by philosophers of quite different persuasions – it is often mentioned in passing that we not only perceive objects or individuals but also perceive events. According to M. G. F. Martin, “Some of the objects of perception – the concrete individuals, their properties, the events these partake in – are constituents of the experience” (2004, 39). Similarly, Tim Crane says, “what is represented in experience are objects, properties and events, in what might loosely be called a ‘manifold’, but which does not have the structure of judgeable content” (2011, 96). Despite philosophers’ willingness to list events among the

entities we perceive, however, individuals or objects have been the predominant focus of theorizing in the philosophy of perception. Casey O'Callaghan reports such a common attitude in his paper: "Given the prominence of objects in visual perception, it is tempting to think that all perceiving concerns objects, their features, and their arrangement" (2008, 803–4).

One might think that there are good methodological reasons to uphold the primacy of individuals in the philosophical study of perception, but such primacy is rarely explicitly defended; it is more like a guiding assumption that is tacitly and widely accepted. An example of this attitude can be found in the work of Barry Stroud. According to him, many conditions must be satisfied for subjects to be "perceptually aware of a fact of the objective world that they thereby know to be so" (2015, 392). The necessary conditions for such perceptual knowledge include the abilities "to be aware of an object and single it out perceptually, to discriminate it from its background or surroundings, to have our attention drawn to it, perhaps to track it if it is moving", among others (2015, 392). But suppose I look out of the window while on an airplane. Since the plane is flying through clouds, I see nothing but moving clouds. Isn't my experience sufficient for me to acquire perceptual awareness, or even perceptual knowledge, of an objective fact about the world? There is certainly such a fact: that there are clouds outside the window. However, it is completely unclear that I have singled out anything, segregated any perceived scene into a figure and a background, or tracked the movement of anything. All I see are clouds that are more or less homogeneous. Though I do see motion, I am unable to identify any sharp spatial boundary that separates two clouds, let alone track the movement of one particular cloud. It then becomes clear that Stroud's argument is based on an implicit assumption: to perceive is to



perceive an individual. Such an assumption seems so natural that it barely requires to be made explicit; there is even less need to defend it. Nevertheless, unless we are ready to reduce all forms of perception to the perception of individuals, such an assumption cannot be made without arguments.

But maybe I am exaggerating things. There has been growing interest in time perception, and perhaps this trend is effective in remedying the lack of attention to event perception. That is unfortunately not true, as philosophers working on the topic of temporal awareness typically attempt to capture something that is common to the perception of events and the perception of individuals. According to Ian Phillips, the datum to be explained by a theory of time perception is that “there are cases... in which one hears or perceives in such a way that one is able to attend to a structure of notes, events or event parts which occupy a temporal interval” (2008, 178). This remark may be read as saying that the main issue about time perception is how we can experience events. But presumably Phillips wouldn’t deny that his theory also applies to cases in which a pumpkin is perceived to remain identical over a short duration of time or a traffic light simply stays red – here pumpkins and traffic lights are obviously individuals. How we can experience these is no less an issue about time perception than, say, how I can experience a laser show, in which no individuals whatsoever are perceived. If a theory of temporal awareness aims to account for both kinds of cases, it cannot tell us what is special about one of them.

It is therefore my contention that philosophers have neglected to give the issue of event perception the attention it deserves. Two notable exceptions are Matthew Soteriou (2010) and Thomas Crowther (2014). Despite their interest in the issue, however, their main concerns are different from mine. On the one hand, Soteriou’s goal is to motivate the

relational view by way of arguing against the representational view. According to him, the representational view has difficulties explaining how one can experience two different phases of an event in a single experience that exemplifies a single content, but the relational view faces no such difficulties (Soteriou 2010, 231–32; 234–36). On the other hand, Crowther’s paper aims to explain how one can see an event, instead of a fragment of the event, even if one is not acquainted with every temporal part of the event (Crowther 2014, 443–45). If I see a layup in a basketball game now, I do not see every temporal part of the basketball game at this moment. Regardless of this, Crowther intends to show that I can still be said to perceive the game itself, not just the fragment of the game that contains the layup. The issues discussed by Soteriou and Crowther are intriguing and important. But their concerns are different from mine, as is demonstrated by my emphasis on the Central Thought Experiment. The questions raised by the thought experiment remain to be addressed, and the theory I am going to propose does exactly that.

## 5.2 Contrast Case 1: The Experience of Individuals

I claim that the perception of events can be distinguished from the perception of individuals on phenomenological grounds. To defend this claim, let’s now consider the way objects are represented in a perceptual experience of individuals. Roberto Casati (2015) discusses various theories of the conditions on which an entity appears as an object of vision, and two of the theories are particularly relevant in the present context. The first theory is based on the work of Palmer and Rock (1994). According to them, an entity is perceived as a basic unit in one’s perceptual environment if it has *connected* and *uniform* parts, and Casati further argues that the connection must be *maximal* (Casati 2015, 394–96). For

example, whereas any of the individual stars and stripes on the U.S. flag meets all three conditions, the total white region on the flag fails to satisfy the condition of maximal connection. The region is indeed uniformly white, and many of its parts are connected; despite these, not *every* pair of qualitatively uniform parts are connected – the stars and the white stripes are not connected, say. The total white region thus fails to appear as a basic unit. In Casati’s terminology, the entities that are perceptually singled out as basic units are called “individuals” or “entry objects” (2015, 396). Given my terminological choice to use the term “object” to denote any perceptually represented entity, I shall call Casati’s entry objects “entry individuals”. Objects that are composed of entry individuals will be called “complex individuals”.

Another theory discussed by Casati, which is proposed by Elizabeth Spelke (1990), is also of interest here. According to Spelke, the computational process correlated with object perception proceeds in a few stages. First, given a surface and a time interval, the visual system keeps track of the location of each surface point over the entire time interval (Spelke 1990, 48). The resulting representations of surfaces provide the materials for further processing. Second, the visual system carves the represented surfaces into physical bodies according to two principles: two surface points are represented as belonging to distinct bodies *if* they are not connected by any continuous path (the *cohesion* principle), and *only if* so (the *boundedness* principle) (1990, 49). Third, if a cohesive and bounded body is moving, its motion is represented according to two additional principles. Whenever possible, the visual system represents the motion as rigid (the principle of *rigidity*) and independent of the motion of any other body (the principle of *no action at a distance*) (1990, 50). For example, consider two basketballs moving away from me in the same direction and at the same speed.

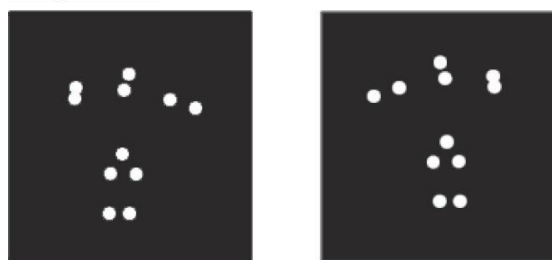
My visual system will not represent the basketballs as shrinking, which is a kind of non-rigid motion; in addition, my visual system will not represent the basketballs as connected by an invisible thread, which would mean that the motion of each basketball is constrained by that of the other.

These two theories suggest that an entry individual is an entity whose parts appear to be qualitatively uniform, maximally connected, disconnected from any other visible surface, and moving in a coordinated way that is characteristic of rigid bodies and independent of any other visible instances of motion. It is clear that perceived events typically fail to exemplify these features. Suppose I am standing by the sea, watching waves break and water wash up on the beach. What I see appears to be a continuous event. However, though parts of the event can be identified, those parts are hardly uniform. I may at once see swash on the beach and the crest of a wave, for example. If the wave is breaking, its crest may appear white; if not, it may appear blue. The swash on the beach, by contrast, appears to have mixed colors. Some portions of it may look transparent; others appear to have the color of sand because of the sand submerged in water; still others may look gold if the water is reflecting sunlight. Not only does the event violate the condition of qualitative uniformity; it also violates the condition of rigid motion. However the parts of the event are individuated, shrinking and expanding are among their common features. The situation is aggravated by the fact that there is no straightforward way to perceptually identify the spatial boundary between the event and the other visible surfaces. The location of the swash is constantly changing, and waves break after they are formed. As a result, no part of the event appears as a cohesive and bounded body capable of undergoing rigid motion.

The experience just described is by no means rare. Consider what happens when one sees the moving fog in a forest, the shadows of leaves in the wind, changes in the lighting conditions of a room, and the great wildebeest migration in Africa (when seen from afar). All these experiences fail to represent an object that features an easily noticeable spatial boundary and apparently independent motion. However, the reason is not necessarily that the kinds of entities perceived in these experiences cannot be represented that way. The shadow left by a moving tennis ball appears to have an easily noticeable boundary. If one sees the shadow but not the ball, then the motion of the shadow also appears unaffected by any other moving entity in the perceived scene. The observation suggests that the distinction between the experiences of individuals and events is primarily a distinction between representational contents. One may represent a motion *event* of which a shadow is a participant or a shadow-looking *individual* whose properties include motion. Whether the *objects* of these experiences are the same *entity* is a partially metaphysical issue that cannot be settled on phenomenological grounds alone. I will set aside this issue for now and focus on a more urgent question: what exactly is the distinctive phenomenological feature that sets the perception of events apart from the perception of individuals?

Here's my proposal: while the perception of events partly consists of the awareness of temporal boundaries, the perception of individuals is not.<sup>72</sup> Though this answer is not particularly striking, I think it makes good phenomenological sense. One way to highlight the difference is to consider point-light displays. Point-

Biological motion



**Figure 2. Point-light displays: an example**  
 Louisa Miller, Hannah C. Agnew, and Karin S. Pilz., *Fig. 1*. From Miller et al. (2018).  
 doi: 10.1016/j.visres.2017.08.004  
 [CC BY 4.0,  
<http://creativecommons.org/licenses/by/4.0/>]

light displays are images of the kind shown in Figure 2, taken from Miller et al. (2018). The two images here depict tennis players about to serve. Now imagine that you are shown a video clip consisting of such images and then given two different tasks. In Task 1, you have to determine what kind of organism is depicted by the video. In Task 2, you have to determine what kind of activity is depicted by the video. What would you look out for in carrying out each task, and how would your experience in one of the scenarios differ from your experience in the other?

I think the following is likely to happen. In carrying out Task 2, one would pay particular attention to the moment when the depicted activity is finished. One would focus on identifying the point at which a seemingly complete pattern of activity has been observed

---

<sup>72</sup> Here the term “temporal boundary” is to be understood intuitively; it refers to anything that intervenes between two temporal phenomena. Thus noon is the temporal boundary between the morning and the afternoon; similarly, Barack Obama’s presidential inauguration was the temporal boundary between his state of being the president-elect and his presidency. This intuitive notion of a temporal boundary is much broader than Crowther’s technical notion of a temporal boundary, discussed in section 3.2. To refer to the kind of temporal boundary Crowther has in mind, I shall use the term “completing boundary”; see Section 5.3 below.

and the same pattern begins to be repeated. In short, one would look out for the boundary of the event in time. This piece of information is essential to the successful completion of Task 2. While the images in Figure 2 could be followed by images that depict different phases of a tennis serve, they could also be followed by images that depict a baseball pitch. If an observer mistakenly attributes a temporal boundary to the event at a moment when the event is still unfolding, she could misrepresent the event, thereby failing Task 2. But none of these matters for Task 1. In carrying out Task 1, one would be attentive to the way the points are spatially configured at each moment – whether the figure formed by the points looks like a whale, a monkey or a human being, etc. One would also track the displacement of the points over time and observe whether the spatial configuration of the points is changing. However, one would *not* try to discern the temporal parts of the depicted object, nor would one deploy one’s attention in such a way as to make oneself alert to any repeated pattern. One simply wouldn’t attempt to identify the point in time at which the object appears to “terminate” in any sense.<sup>73</sup> To be sure, one could go about identifying the organism by first observing the type of activity that is taking place. In that case, one’s experience in Task 1 would be much more similar to one’s experience in Task 2. What follows from such a possibility, however, is that the perception of events may facilitate the perception of individuals. My point still stands: the awareness of temporal boundaries is the

---

<sup>73</sup> Casati and Varzi have made a related point about the metaphysical difference between individuals and events: whereas individuals feature easily discernible spatial properties, events are characterized by salient temporal properties (Casati and Varzi 1999, 169–70). My concern here is the implications such a contrast has for our experiences of individuals and events.

main phenomenological feature that distinguishes the perception of events from the perception of individuals.

The question that then arises is how exactly such awareness should be understood. Is it really perceptual? Can the Husserlian dual-component view defended in the previous chapter accommodate such awareness? These questions will be addressed in the following sections. For now, I would like to emphasize that the awareness of temporal boundaries isn't anything mysterious. Though I am not aware of existing work that explicitly takes such awareness to be the phenomenological criterion that sets the perception of events apart from the perception of individuals, such awareness has been extensively studied in the psychology of event individuation. As noted by the psychologists Tanya Sharon and Karen Wynn, event individuation is a pervasive phenomenon crucial to many everyday activities. They argue that one must be able to parse an event into units if one is to learn the steps necessary for carrying out an action, separate the cause and effect in a represented causal chain, identify the meaningful units of others' actions, and refer to the events others refer to (1998, 357). Even if one is reluctant to accept the view that we can differentiate between the experiences of events and individuals by appealing to the kind of awareness that makes event individuation possible, there is nothing controversial about the claim that there *is* such awareness.

### **5.3 Contrast Case 2: The Experience of Time**

It is my contention that the significance of event perception is underappreciated in philosophy; too much theorizing rests on the implicit assumption that individuals are the paradigm of perceptual objects. One may think that this is an overstatement, however. The



experience of time has always been a central concern of philosophers. Since the experience of time obviously involves the awareness of temporal boundaries, it is tempting to think that all the phenomenological features of event perception are derivative of those of time perception. A potential point of departure for such a view is the common observation that the temporal location of a perceived event typically appears to extend beyond the now. For example, Joel Smith claims that to perceive an event, we must be aware of what has just happened: “For us to have an experience of an event *as an event*... there must be, at any one moment, some sort of awareness of the just past phases of the event simultaneous with the awareness of the present phase” (2016, 93). This suggests a way to conceptualize event perception, which I shall call the theory of *event perception as time perception* (EAT):

(EAT, First Pass)

Let  $v$  be a visual experience and  $o$  the object of  $v$ . Then  $v$  counts as an instance of event perception if and only if:

- (1) in undergoing  $v$ , the subject is aware of the temporal boundary between the past and the present;
- (2) the temporal location of  $o$  appears to overlap both the past and the present.

EAT appears to capture Smith’s idea quite well. However, Smith is prudent in refraining from taking his idea to give us a sufficient condition of event perception. The present formulation of EAT leaves much to be desired. In what follows, I will refine EAT in an attempt to turn it into a plausible account of event perception. As can be expected, however, the conclusion reached in this section will be that EAT alone cannot satisfactorily account for event perception.

The first issue is how exactly the present should be understood in the framework of EAT. Should it be equated with a particular *instant*? That cannot be right. When I point to a bus and say, “there is a bus moving towards me now”, the referent of the indexical “now” encompasses more than a single instant. No moving can possibly occur within a single instant. What “now” refers to is presumably what William James famously describes as the *specious present*: “the original paragon and prototype of all conceived times is the specious present, the short duration of which we are immediately and incessantly sensible” (1890, 1:631, emphasis omitted). The “now” in the sentence under consideration does not denote an instant that can be singled out with some physical measure independently of my mental states; rather, it denotes an interval. An instant belongs to the interval insofar as it is the temporal location of an object that *appears to be in the present*. But what would it mean to say that an object appears to be in the present? Does it simply mean that the object figures in the experience that one is undergoing? Barry Dainton’s view could be interpreted this way. According to him, “An experience of succession involves a temporal spread of contents being presented *together* in consciousness, albeit in the form of a perceived succession rather than simultaneously”; moreover, “Contents which are apprehended as unified in this way belong to a single specious present” (2018, sec. 1.2, para. 3). Take Edith’s experience of a stone being thrown into water. Since Edith doesn’t perceive the resulting ripple as emerging out of the blue, it must be the case that she is still aware of the falling of the stone. It follows that the ripple and the falling are “presented together in consciousness”, so they both appear to be in the present.

Such a view is equivocal at best and phenomenologically untenable at worst. Suppose that after undergoing a visual experience of a green cantaloupe over interval  $I_1$ , I keep my

eyes closed over  $I_2$  and imagine the way the cantaloupe would look if it were to be painted black. To do so, I must continue to represent the green cantaloupe over  $I_2$ , or at least over the earliest portion of  $I_2$ . In fact, on the assumption that  $I_2$  immediately succeeds  $I_1$ , the green cantaloupe is putatively represented over the initial portion of  $I_2$  no matter what I end up doing over  $I_2$  – this is the case insofar as my short-term memory is functioning normally. We can then conclude that according to the view under consideration, there is a time at which the green cantaloupe and the black cantaloupe both appear to be in the specious present. Since they are not experienced as simultaneous, we are compelled to say that I thereby undergo an experience of succession. This is nonetheless a weird conclusion. Even if both my experience of the cantaloupe and Edith’s experience of the stone can be regarded as experiences of succession, they are obviously quite different. Whereas the successive contents of Edith’s experience represent the temporal relation *between the objects themselves*, it is unclear that the same can be said of the successive contents of my experience. My experience tells me very little about how the green cantaloupe is temporally related to the black cantaloupe. The view is thus equivocal between two ways of understanding the phenomenon of co-occurrence in the specious present: such co-occurrence can be construed as a relation between the *contents* of experiences and the *objects* of experiences, but the view fails to distinguish the two.

Recall that our goal is to describe the way the temporal boundaries between events are perceptually represented. To do so, we should focus on the temporal relations between perceptual objects, not those between perceptual contents. This should be reflected in our conception of the specious present: even if an object is represented by the experience one is undergoing, it doesn’t follow that the object appears to be in the specious present. Here’s a

reasonable additional requirement: the object also has to be represented as *simultaneous with* the experience one is undergoing. Since the green cantaloupe is hardly represented as existing simultaneously with the black cantaloupe, the additional requirement rules out the content-based understanding of the specious present. But how should we accommodate experiences of succession? We could divide the experience one is undergoing into different temporal parts:

(Experience of Succession)

Let  $v$  be a visual experience a subject is undergoing and  $o$  its object. Then  $v$  is an experience of succession if and only if there are  $v_1$ ,  $v_2$ ,  $o_1$  and  $o_2$  such that:

- (1)  $v_1$  and  $v_2$  are among the temporal parts of  $v$ ;
- (2)  $o_1$  and  $o_2$  are among the temporal parts of  $o$ ;
- (3) both  $o_1$  and  $o_2$  are represented by  $v_2$ ;
- (4)  $v_2$  represents  $o_2$  as simultaneous with  $v_2$ ;
- (5)  $v_2$  represents  $o_1$  as simultaneous with  $v_1$ ;
- (6) the subject undergoes  $v_2$  immediately after she undergoes  $v_1$ .

Whenever condition (3) is satisfied, I shall also say that  $o$  is represented as occurring in the specious present.

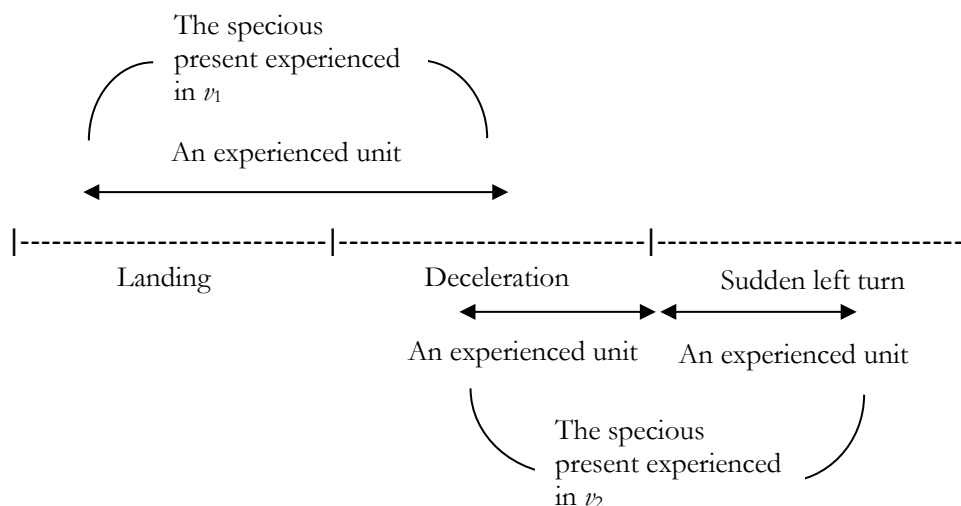
This criterion tells us that Edith's experience is an experience of succession insofar as the later part of her experience represents the falling of the stone as simultaneous with the earlier part of her experience, on the one hand, and represents the ripple as simultaneous with the later part of her experience, on the other. If this is the case, then the entire event appears to Edith as occurring in the specious present.

Note that if we are asked to give an example of an ordinary experience of an event, Edith's experience would work perfectly well. This suggests that cases of event perception just are experiences of succession:

(EAT, Second Pass)

Let  $v$  be a visual experience and  $o$  the object of  $v$ . Then  $v$  counts as an instance of event perception if and only if it is an experience of succession.

Any experience that meets the condition features "some sort of awareness of the just past phases of the event simultaneous with the awareness of the present phase" (J. Smith 2016, 93). Does this version of EAT give us what we want? No, it doesn't, unfortunately. Consider what happens when I change TV channels. I had been watching the live broadcast of a concert featuring many musicians until Krystian Zimerman finished his performance. Then I immediately switched to the live broadcast of a tennis game and saw Roger Federer serving. In this scenario, Federer's serve was perceived to succeed Zimerman's performance, so they were both perceived to occur in the specious present. However, the two events were not thereby perceived to constitute a single unified event. Contrast this case with one in which I saw Novak Djokovic return the ball after Federer's serve – in the latter case, the two events were perceived to form a single unified event. The difference between these two cases, I submit, is a phenomenological one. This suggests that perceiving an event involves perceiving a kind of unity that holds the temporal parts of the event together, and such unity cannot be simply equated with co-occurrence in the specious present. If so, the newer version of EAT still fails to capture the phenomenon I am interested in.



**Figure 3. The experienced instances of the specious present and the experienced units of the perceived event.**

While parts of the landing and the deceleration are experienced as making up a single subevent, parts of the deceleration and the sudden left turn are experienced as distinct subevents. Note that this diagram is an abstraction; there could be more experiences intervening between  $v_1$  and  $v_2$ .

Here's another example of the kind of perceived unity characteristic of events (see Figure 3). Anton is watching the landing of an airplane. After the plane lands, it decelerates and then makes a sudden left turn. Let's focus on two of Anton's visual experiences,  $v_1$  and  $v_2$ . In  $v_1$ , the final part of the landing event appears to be succeeded by the initial part of the deceleration event, where such succession appears to occur in the specious present. In  $v_2$ , the final part of the deceleration event appears to be succeeded by the initial part of the sudden left turn, where such succession again appears to occur in the specious present. I maintain that there is an important phenomenological difference between  $v_1$  and  $v_2$ , a difference similar to the one described in the TV watching scenarios. The events that are the objects of  $v_1$ , i.e. the landing and the deceleration, are perceived to form a single unit. By contrast, the events that are the objects of  $v_2$ , i.e. the deceleration and the left turn, are not so perceived.

There is an intuitive explanation of the difference: in typical cases, a landing event is followed by a deceleration event, but the connection between a deceleration event and a sudden left turn is much more arbitrary. After deceleration, a plane may turn right, turn left, go straight, stop, or accelerate and take off again. While a deceleration event may be perceived as the continuation of a landing event, a sudden left turn would appear as a new subevent that jointly makes up a larger event with the earlier subevent of landing but is not a phase of the latter. If so, however, it would seem that EAT doesn't have the resources to account for the phenomenological difference between  $v_1$  and  $v_2$ . Insofar as the events represented by  $v_1$  can be said to co-occur in the specious present, the same can be said of the events represented by  $v_2$ ; this is the case despite the fact that the events represented by  $v_2$  are experienced as distinct subevents. To put it differently, insofar as  $v_1$  can be said to represent the temporal parts of its object as simultaneous with its own temporal parts, the same can be said of  $v_2$ . Since such simultaneity is the sole concern of EAT, EAT gives us little to go on if our goal is to understand how the experienced event appears differently in  $v_1$  and in  $v_2$ .

This point can be reinforced by considering the work done by Newtonson (1973), briefly mentioned at the beginning of this chapter. He conducted a series of experiments to show that the number of temporal boundaries a subject perceives in an event is affected by the degree to which the subject perceives the event to unfold in a predictable manner. In one experiment, subjects were divided into two groups and shown slightly different video clips. The video clip viewed by the control group depicted the process in which an actor assembled the model of a molecule. The video clip viewed by the experimental group was identical except that an additional sequence of actions was inserted. After spending two minutes assembling the model, "the actor turned to his right, removed his right shoe and

sock, put the sock in the shoe and placed them on the table to his left, bent down again, and rolled his left pants leg up to his knee” (Newtson 1973, 35). This weird series of actions took place over 30 seconds. After that, the actor resumed assembling the model in exactly the same way he continued to assemble the model in the video watched by the control group. The subjects in both groups were asked to record the instants at which a new event was perceived to occur. The result was that the subjects in the experimental group began to identify more units in the depicted event after seeing the inserted sequence; in Newtson’s terminology, their average rate of unitization was 3.37 units per minute, as contrasted with the control group’s rate of 2.67 units per minute (1973, 35). This result suggests that the participants perceive more temporal boundaries that separate subevents when an unfamiliar event is perceived.

Now consider the experience that a subject in the experimental group underwent after she had seen the unexpected of turn of events in the video. We ask: how exactly did her experience differ from the corresponding experience had by a subject in the control group, such that she perceived more temporal boundaries between subevents in the same depicted event? The difference cannot be that she began to perceive time differently; it is not the case that she somehow began to experience a shorter specious present. If the subject had always been able to perceptually represent two 500-msec events as successively happening in the specious present, it is implausible to assume that such an ability was mysteriously lost after she saw the weird series of actions. Hence, though the subject’s experience of time remained the same, her experience of events had changed. As things stand, nothing in EAT captures the difference between the experiences had by the subjects in the experimental and control groups.



What emerges from these cases is that insofar as our goal is to describe the distinctive phenomenological features of event perception, the central issue shouldn't be whether an experienced event is perceived to occur in the specious present. What really matters is how a subject perceptually identifies the subevents or units of the perceived event. Consider again Anton's experience of the aircraft. Let's say that at  $t$ , the aircraft had just begun to make the sudden left turn. While we could ask whether the left turn and the deceleration event preceding it were perceptually represented as successively occurring in the specious present, the more important question given our concern is whether they were represented as jointly constituting a single event. If the two events were perceived as two different phases of one unified event – call this unified event “Taxiing” – then Anton's experience at  $t$  represented Taxiing as ongoing. If, by contrast, Anton perceived little continuity between the deceleration and the left turn, then his experience at  $t$  represented the deceleration as having ended and being succeeded by a new event, i.e. the left turn. In such a scenario, Taxiing played no role in Anton's experience. Taking this approach to the issue of event perception, we can formulate the key phenomenological questions to be addressed in this chapter as follows:

(Central Questions)

- (1) What is it like to see an event as ongoing?
- (2) What is it like to see an event as ending?

To defend the idea that event perception is a species of time perception, we have to show that this pair of questions can be answered by invoking nothing but the phenomenological features of our temporal awareness. I have argued that the two versions of EAT examined

above are inadequate theories of event perception. Now that we have a clearer sense of the kind of answers we are looking for, however, maybe we can formulate yet another version of EAT that better addresses the Central Questions:

(EAT, Third Pass)

Let  $v$  be a visual experience and  $o$  the object of  $v$ . Then  $v$  counts as an instance of event perception if and only if:

- (1) EITHER  $v$  represents  $o$  as ongoing, which is the case if and only if: for any  $x$ , insofar as  $x$  is a temporal part of  $o$  that appears to be located in the specious present,  $x$  doesn't appear as the final temporal part of  $o$ ;
- (2) OR  $v$  represents  $o$  as ending, which is the case if and only if: there is some  $y$  such that  $y$  is a temporal part of  $o$  that appears to be located in the specious present and  $y$  appears as the final temporal part of  $o$ .

To analyze Anton's experience at  $t$  in terms of this version of EAT, we have to determine whether Anton's experience represented any of its objects as the final temporal part of some event. If, for example, Taxiing was an object of Anton's experience but nothing represented by the experience was represented as the final temporal part of Taxiing, then EAT entails that Anton perceived Taxiing as ongoing.

This version of EAT is an improvement over the first two versions, but it has a major flaw. It doesn't tell us what it takes for an experience to represent an object as the final temporal part of an event. Given an event, the experiences of the event can be divided into two classes: those that represent the event as the final temporal part of some other event and those that do not. How exactly do the experiences in the first class differ from those in the

second? I don't see how we can answer this question by focusing exclusively on the features of temporal awareness. Therefore, while this last version of EAT may be useful in clarifying the way time perception interacts with event perception, it is not in itself an appropriate account of the latter. Though time perception and event perception both require some sort of awareness of temporal boundaries, they represent different kinds of temporal boundaries. In our awareness of time, we experience the boundaries of the specious present. In our awareness of events, by contrast, we experience what I call the *completing boundaries* of events. Before we turn to the idea of a completing boundary, however, it is worth emphasizing that I have no intention to claim that event perception can occur independently of time perception. The only thing I insist on is that the latter isn't sufficient for the former.

The notion of a completing boundary is derived from Crowther's technical notion of a temporal boundary. According to the theory proposed by Crowther (2011),<sup>74</sup> no event can occur without being completed by the kind of temporal boundary appropriate to its nature. A watering of plants cannot occur without some plant's being watered; a plant's being watered is thus the temporal boundary that is required by, and completes, a watering of plants. Crowther holds that such temporal boundaries should be understood in terms of temporal stuff: a moment in time is the location of a temporal boundary if certain temporal stuff comes into being or ceases to be at that moment. I do not subscribe to Crowther's distinction between temporal stuff and temporal particulars, but I find the idea of event completion compelling. Hence, instead of adopting the definition of a completing boundary based on temporal stuff, I will take the idea to be simply a generalization of what Vendler

---

<sup>74</sup> The theory is discussed in section 3.2 of this dissertation.

calls the climax of an action, i.e. a point that “has to be reached if the action is to be what it is claimed to be” (Vendler 1957, 145).<sup>75</sup> Vendler’s idea is now widely accepted, and it is customary for linguists to draw the distinction between *telic* and *atelic* events. Here’s how Kate Kearns draws the distinction: “The property of **telicity**... is the property of having a natural finishing point... Any event which does not have a natural finishing point is **atelic**” (Kearns 2011, 157). But as Crowther notes, an atelic event just is an event that can be completed by any kind of ending point (Crowther 2011, 24). Meditating is an atelic event, and one can stop meditating by standing up, turning on the TV or beginning to eat. Doing any of these completes the meditating event, even though the event doesn’t have a climax. I will therefore take the idea of a completing boundary to be applicable to both telic and atelic events: something is a completing boundary of an event insofar as it is capable of individuating the event in a way compatible with the event’s nature. That is, given events  $e$  and  $f$ ,  $f$  is a completing boundary of  $e$  if and only if ending with  $f$  counts as a natural course of  $e$ . Note that I speak of *a* completing boundary instead of *the* completing boundary of an event, because there may be multiple final phases an event could reach when it runs its normal course. This is especially the case for atelic events. Of course, any event that has already occurred in the actual world has no more than one completing boundary.

So construed, the notion of a completing boundary is closely related to the notion of an event boundary that is widely adopted in the psychology of event individuation. However, the term “event boundary” is often used by psychologists in such a way as to make its metaphysical import obscure. This happens because psychologists do not typically work with

---

<sup>75</sup> Again, see section 3.2 for discussion.

a metaphysically nuanced conception of events. For example, in one of their studies, Jeffrey Zacks and Barbara Tversky claim to focus on “the following archetype for an event: a segment of time at a given location that is conceived by an observer to have a beginning and an end” (2001, 3). Such a definition of events lumps together the boundaries of a time interval and those of an event; moreover, it seems to suggest that the boundaries of an event are what an observer takes them to be. These ambiguities make the definition inadequate for the present task, i.e. to describe how the way events are represented in perception differs from the way other entities, including time, are represented in perception. Hence, though I will draw on the psychology of event individuation to defend my view, I will work with a view that takes completing boundaries to be integral to events, which has a more explicit realist character than a view like Zacks and Tversky’s. The experiences of events are phenomenologically distinguishable from those of individuals and time precisely because they have contents that represent completing boundaries. In what follows, I will clarify and defend this claim.

## **5.4 The Experience of Events**

### *5.4.1 Experiencing Completing Boundaries*

If focusing on our awareness of the specious present alone isn’t enough for us to understand the way the completing boundaries of events are experienced in conscious visual perception, what else should we consider? The first step, I submit, is to invoke the expectational contents described by the Husserlian dual-component view (HDC). The reason is that whether the completing boundary of an event appears to have occurred

typically depends on what one expects to happen at the following moment. Suppose there is water in front of Franz and Franz has just dropped his phone. At this point, does the submergence of the phone in water appear to Franz as a completing boundary of the phone's falling? Not if Franz perceives the water to be just a puddle. In that case, Franz expects the phone to remain visible at the following moment; the possibility that the phone becomes submerged in water wouldn't even strike Franz, let alone figure in the content of Franz's experience. Hence, what would appear as a completing boundary of the phone's falling depends on what Franz expects to happen when the phone actually touches the ground. To adequately characterize Franz's visual experience of the phone's falling, we must pay attention to the events represented by the experience's expectational contents.

Expectational contents carry information about what is likely to occur in the immediate future. Likelihood is a modal notion. Therefore, if we are to describe the expectational contents of event-representing experiences, a convenient framework is that of possible world semantics. It is nonetheless important to note that I am not committed to the idea that possible worlds are the *contents* of experiences.<sup>76</sup> The only kinds of contents relevant to my analysis here are those posited by HDC, i.e. factual and expectational contents. Possible worlds are only invoked to specify the *objects* of experiences; the framework serves the sole purpose of pinpointing the future event that is represented as what will probably succeed the currently perceived event. It follows that I will have little to add to what has

---

<sup>76</sup> According to Michael Tye, "a natural suggestion is that the content of a visual experience is simply a set of possible worlds, namely the set of worlds at which the experience is accurate" (2014, 300). I do not endorse such a view. For my take on the relation between possible worlds, perceptual contents and perceptual objects, see section 5.4.5.

already been said in the previous chapter about the nature of expectational contents themselves – the proposal presented here won't shed further light on, say, whether such contents have a compositional semantics or what their atomic components are. The analysis developed here presupposes HDC rather than alters it. To be sure, I will supply additional details as we go to turn HDC into an appropriate account of event perception, but those details should be compatible with everything I said in the previous chapter.

Here's how I think we should use possible worlds to capture the expectational contents of event-representing experiences.<sup>77</sup> I call my proposal the *boundary representation view* about event perception (BRV):

(BRV, First Pass)

Let  $v$  be a visual experience whose subject is  $s$  and whose object is event  $e$ .

- (1)  $v$  represents  $e$  as ongoing if and only if in every possible world expectable for  $s$ , there is an event  $f$  of which the following are true:
- (a)  $f$  occurs immediately after  $e$ ;
  - (b)  $f$  jointly makes up a larger event with  $e$ .

---

<sup>77</sup> The idea that some features of perceptual contents can be captured with possible worlds is not new. An early attempt to do so is Hintikka (1969). Within the phenomenological tradition, Smith and McIntyre (1982) and Banick (2017) have responded to Hintikka's work. Outside the phenomenological tradition, possible worlds appear in the theory of perception proposed by Egan (2010). What makes my focus different from all these scholars' is that the possible worlds invoked in my account are future possibilities: they are what the present actual world might become in the immediate future. By contrast, the possible worlds that figure in the above accounts are either alternatives to the present actual world or worlds defined independently of times. The cases that my proposal is designed to address are therefore somewhat different from those that the above accounts focus on.

(2)  $v$  represents  $e$  as ending if and only if in every possible world expectable for  $s$ , there isn't any event that satisfies both (1)(a) and (1)(b).

If (1) is the case,  $s$  experiences the completing boundary of  $e$  as yet to occur. If (2) is the case, the currently perceived temporal part of  $e$  appears as the completing boundary of  $e$ .<sup>78, 79</sup>

Let's analyze the aforementioned case of Anton's experience in terms of BRV. Recall that Anton perceived the successive occurrence of an aircraft's landing and deceleration. What must have been the case for Anton to see the landing as ongoing? Obviously, the factual contents of Anton's experience had to represent the landing. According to BRV, it was also necessary for the expectational contents of his experience to represent an event that would happen immediately after the landing and jointly make up a complex event with the landing. This condition was presumably satisfied because Anton had seen enough aircraft landings in

---

<sup>78</sup> An alternative is to say that the latest temporal part in the specious present appears as the completing boundary of  $e$ .

<sup>79</sup> At this point, the reader might begin to think that there is something weird about the way the Central Questions and BRV are formulated. Why focus on the ending of an event rather than *both* the ending and beginning of an event? There are two reasons for this. First, in many cases, the perception of an event's beginning simply consists in the fulfillment of one's prior expectation that the earlier event is ending. Suppose I saw an accelerating aircraft at  $t$  and expected the acceleration to end momentarily with the takeoff of the aircraft. My expectation was fulfilled at the moment immediately after  $t$ . This seems sufficient for me to experience the beginning of another event, i.e. the aircraft's ascent – given that the takeoff was perceived as the ending of an event, what appeared immediately after it was perceived as the initial phase of a new event. Second, in those cases where one perceives the beginning of an event without having expected it, one's experience can be analyzed in a way largely analogous to the analysis given in BRV. The main difference is that instead of expectations, we need to invoke one's awareness of the events in the immediate past. Suppose the aircraft in our example took off earlier than I expected, so I had not formed the expectation that the acceleration process was about to end. However, when I saw the takeoff, I was aware (1) that the aircraft had been accelerating, and (2) that the takeoff was not a part of the acceleration process. I thus perceived the takeoff as initiating an event different from the one that just happened. In short, if my analysis here is correct, then it shouldn't be too difficult to accommodate our experiences of event beginnings within the framework of BRV. To simplify the arguments, therefore, I will focus on how we perceive the endings of events.



the past. Given Anton's past experience, upon seeing the aircraft land, Anton automatically expected to see a process of deceleration – such a process appeared to Anton as the only natural way for the event to continue. As a result, any possible world in which such a process failed to occur would defy Anton's expectation, so both conditions (1)(a) and (1)(b) were met. BRV then tells us that Anton perceived the completing boundary of the landing as yet to occur, so the landing was perceived as ongoing.<sup>80</sup>

As things stand, the conditions specified by BRV are not precise enough. An issue that has to be addressed immediately is the criterion of an expectable world. What conditions must be met for a world to belong to the set of expectable worlds? It cannot be right to say that the set includes every world in which I undergo an experience that is possible for me to undergo after experiencing what I am experiencing now. After seeing a landing event, Anton could see a deceleration event, a police chase or the eruption of a nearby volcano. The last two events are rather unlikely, but they are certainly metaphysically possible – or even epistemically possible. Despite this, for the purpose of singling out the objects represented by the expectational contents of Anton's experience, it would be wrong to include those worlds. As I will try to show, choosing the expectable worlds appropriately is no easy task. One of the complications results from the fact that there may be substantial

---

<sup>80</sup> To make the claim more precise, we should say that what was perceived as ongoing was the complex event jointly made up by the landing and the deceleration. This requires us to change condition (1) in BRV: the phrase "*v* represents *e* as ongoing" should be replaced by "*v* represents an event of which *e* is a temporal part as ongoing". Since the former is simpler and more easily readable, I will stick to the former. It should nevertheless be understood that the simpler phrase is just a convenient expression that abbreviates the more precise version of condition (1) in BRV.

interpersonal differences in the way the same event is experienced – this will be our topic in the upcoming section.

#### 5.4.2 *An Excursion into Natural Language Semantics*

Given any event, the way the event unfolds would be familiar to some people but not to the others. If we want a theory capable of accommodating all these people's experiences, we run into a tricky issue: how should such familiarity be reflected in our choice of the relevant possible worlds? On the one hand, the worlds expectable for an expert vastly outnumber the worlds expectable for an amateur. Suppose I am invited to watch a sumo competition for the first time. When the competition is about to begin, I see two men in unusual clothing crouching in a ring. I do have some expectations about what will happen: in light of my past experience with other forms of sports, I expect the men to stay in the ring and keep their distance from the referee. But these pretty much exhaust all my expectations. Since I have never watched a sumo competition, I have no expectations whatsoever about how the men will interact with each other. This is not the case for Johannes, who is a Japanese culture enthusiast that regularly attends such competitions. For Johannes, the very way a sumo wrestler crouches already conveys a lot of information about what the sumo wrestler is about to do. Therefore, in specifying the expectational contents of Johannes' and my perceptual experiences of seeing the beginning of a sumo competition, we have to use different sets of possible worlds. Some worlds contain events that Johannes' experience would, but mine would not, represent as likely to occur at the following moment; those worlds should be excluded when we try to describe the expectational contents of my experience.

On the other hand, there is a sense in which we need to take into consideration more possible worlds when specifying the expectational contents of an amateur's experience. Suppose I woke up in a venue for sumo competitions with no idea how I got there. I saw two men standing in a ring, but I didn't realize that they were about to start wrestling. Depending on what's going on in my mind, I might expect the men to start dancing as part of a religious ritual or to both charge at the referee. These possibilities would not even come across Johannes' mind if he were in my position. It follows that many course of events might appear possible to me but not to Johannes. The possible worlds in which the perceived event unfolds in those ways should be excluded when we try to pin down the events likely to be represented by the expectational contents of Johannes' experience.

If the above reasoning is correct, then the set of worlds expectable for an expert is both larger and smaller than the set of worlds expectable for an amateur that is perceiving the same event. How can that be?<sup>81</sup> The way out of this thorny situation, I submit, is to use more than one set of possible worlds. The work of the linguist Paul Portner provides a good model of how this can be done. Portner defends a semantics of the progressive aspect that builds on an earlier theory proposed by David Dowty (1977, 1979). Portner uses the sentence "John was building a house" to illustrate Dowty's theory, which maintains that the sentence "is true iff something was going on that, if it had proceeded normally, would have been a complete house-building by John" (Portner 1998, 762). Portner's semantics has two

---

<sup>81</sup> This question is to some extent rhetorical. Strictly speaking, when compared with the set of worlds expectable for the amateur, the set of worlds expectable for the expert is not *both* larger and smaller; it is larger in one sense but smaller in another. Still, the question of how these are to be reconciled remains.

major theoretical constructs, which aims to flesh out the idea of a complete event, on the one hand, and that of normal progression, on the other.

Take the sentence “At noon, Drone No.89 was making a delivery to Roman”. How do we identify the truth conditions of this sentence? According to Portner, we have to consider the *modal base*, which is a set of facts that collectively specify the nature of the event described by the verb predicate in the progressive form. As Portner puts it, the modal base is “the set of circumstances relevant to whether e is completed” (1998, 774).<sup>82</sup> In our example, the facts in the modal base collectively guarantee that the ongoing event described by the predicate “was making a delivery to Roman” would turn into a complete event of Drone No.89’s delivery of ordered merchandise to Roman, provided that the event was free from interruption. The modal base is therefore made up of facts concerning Drone No.89 and its activities around noon, such as that the drone was perfectly functional, that it indeed carried some ordered merchandise, and that the drone had already left the storage facility, etc.

In addition to the modal base, we also have to consider the *ordering source*, which “provides a definition of what it is for a given event not to be interrupted” (1998, 777). In our example, the definition would be a conjunction of such conditions as that Drone No.89 wouldn’t get attacked by aggressive birds, that there wouldn’t be malicious attempts to hack into the drone, and that the drone wouldn’t get hit by a car, etc. Once we have singled out the modal base and the ordering source, we follow a two-step process to determine the truth conditions of the sentence. First, we identify the set of worlds compatible with the modal

---

<sup>82</sup> The idea of a modal base is originally proposed by Angelika Kratzer ([1981] 2002). Similarly for the idea of an ordering source, to be discussed shortly.

base. Second, we examine the subsets of the set and choose the subset that best satisfies the ordering source. Then we say that the sentence is true if and only if in all the worlds chosen in the second step, there is an event  $e$  such that Drone No.89's activity at noon was a temporal part of  $e$ , on the one hand, and  $e$  ends up being a successful delivery of ordered merchandise to Roman, on the other (1998, 782).<sup>83</sup> This is how Portner fleshes out Dowty's idea: the modal base captures what an event is and what it will end up being in those worlds where it is allowed to run its course, while the ordering source captures what it takes for the event to actually run its course.

In a sense, what I propose to do is to formulate a theory of event perception analogous to Portner's theory of progressive sentences. More specifically, we could say that the perceptual analogues of the modal base and the ordering source jointly determine the set of expectable worlds relevant for determining the expectational contents of a perceptual experience. Since the set of expectable worlds associated with an experience is changed if either the modal base or the ordering source is changed, we can make sense of the seemingly contradictory claim that the set of worlds expectable for an expert at once contains more worlds and less worlds than the set corresponding to the experience of an amateur watching the same event unfold. Unfortunately, there are at least two obstacles to pursuing such a suggestion literally.

---

<sup>83</sup> This is a simplification; Portner's actual semantics features many subtleties left out here. For example, the ordering source doesn't simply function as a set of necessary conditions; it is supposed to yield an ordering of the worlds compatible with the modal base. However, the main reason Portner's theory is brought up here is that it illustrates the distinction between the modal base and the ordering source; the fine details about the truth conditions of progressive sentences need not concern us here.

First, Portner's semantic theory is concerned with the truth conditions of progressive sentences, but I am interested in the objects of expectational contents. While some may be inclined to assimilate objects to truth conditions, I have no intention of doing so. On my view, even if an experience is completely veridical with respect to an environment, it doesn't follow that everything in the environment is an object of the experience. The distinction is even more important in the case of possible worlds – even if an experience is, say, completely veridical at the actual world at 5 p.m., it doesn't follow that the experience represents everything in the actual world at 5 p.m. It is thus unclear how Portner's account of truth conditions can be translated into an account of perceptual representations. Second, the situation is even direr if we want our theory of event perception to be phenomenologically adequate. We could try to turn Portner's semantic theory into a psychological theory of beliefs and say something like the following: a person's belief about an ongoing event is accurate if and only if the person's beliefs about the modal base and the ordering source associated with the event are accurate. Whether or not a plausible conception of beliefs can be developed along these lines, it seems unlikely for us to come up with a reasonable theory of event perception by taking this route. After all, what would be the *perception* of a modal base or an ordering source? The idea that we can perceive the latter is especially unintelligible. In the case of Drone No.89, what would it even mean to say that one perceives the absence of any hackers trying to derail the delivery attempt or the absence of strong electromagnetic fields that would eventually disable the drone?

Despite these difficulties, it is my belief that the notions of the modal base and the ordering source can be fruitfully put to use in a theory of event perception. It still seems reasonable to say that when an expert and an amateur perceive the same event, the sets of

worlds expectable for them are different; it's just that these worlds cannot be singled out the way the worlds compatible with the modal base of a sentence are singled out. Insofar as we take these worlds to play a role in determining the objects of the experiences without functioning as the objects themselves, the above difficulties can be circumvented.<sup>84</sup> Let's say that an expert's experience of an event has a wider *expectational base* than the experience of an amateur watching the same event unfold. This explains why, in the aforementioned sumo case, the worlds expectable for Johannes are greater in number than the worlds expectable for me. An expectational base is a set of possible worlds; roughly, it is the theoretical counterpart of a set of worlds compatible with a modal base. Besides the difference in the expectational base, there is also the difference between the *salience constraints* imposed by Johannes' experience and mine, where a salience constraint is basically the theoretical counterpart of an ordering source. A salience constraint separates the relevant worlds in the expectational base from the irrelevant ones. The worlds that meet the salience constraint are called *expectationally relevant worlds*. In the sumo case, the worlds that are expectationally relevant for me are greater in number than the worlds expectationally relevant for Johannes, even though the expectational base of Johannes' experience is wider than the expectational base of mine. This explains why my expectations about what the sumo wrestlers are about to do are more indeterminate than Johannes'; many possibilities that appear relevant to me simply fail to appear this way to Johannes.

---

<sup>84</sup> If one takes this route, it is imperative for one to say something about the exact role the worlds play in determining the objects of the experiences. I will return to this issue in section 5.4.5.

With these concepts in hand, we can give a more precise characterization of the aforementioned issue about BRV: given the expectational contents of an event-representing experience, what criteria should we use to identify the set of possible worlds relevant for specifying the objects of these contents? We can now take the issue to be: given an experience of an event, how should we identify the worlds that are expectationally relevant for the person undergoing the experience? The answer is: we should look for the expectational base of the experience and the salience constraint imposed by the experience. Obviously, both concepts must be defined more precisely; that will be the topic of the next two sections. A good way to conceptualize the expectational base, I suggest, is to appeal to what Jeff Yoshimi calls “expectation gradients”. I will therefore turn to Yoshimi’s work now.

#### 5.4.3 *Specifying Expectational Contents: The Expectational Base*

In an attempt to make Husserl’s view on the experience-expectation relation formally precise, Yoshimi proposes what he calls the *expectation rule*, which has the following form:

$$f: K \times V \times B \rightarrow [0, 1]^V \quad (\text{Yoshimi 2016, 18})$$

The meanings of the symbols are as follows:

- $K$  is the set that includes exactly those possible states that could function as a person’s background knowledge: “A state of background knowledge corresponds to all of a person’s unconscious beliefs, attitudes, tendencies, and understandings at a time. These states persist while we sleep and have an impact on what we do when we are awake” (2016, 16).



- $V$  is the set of all sensory impressions that are possible for human vision.<sup>85</sup> Yoshimi illustrates the idea of a sensory impression by comparing the experience of a house and the experience of a façade of the house in a movie set. The component shared by these two experiences is the sensory impression of a house-looking entity (2016, 13–14).
- $B$  is the set of all kinesthetic experiences that could be had when a person moves her body.
- $[0, 1]^V$  is the set of all functions from the set  $V$  to the interval  $[0, 1]$ , i.e. the set of all numbers between 0 and 1. The interval represents degrees of fulfillment or “degrees to which what we see confirms or disconfirms prior expectations” (2016, 15).

As a whole, the expectation rule says that the degree to which a person’s sensory impression at  $t$  would confirm her earlier expectations depends on three aspects of her mind at the moment just before  $t$ : her background knowledge, sensory impression and awareness of her own bodily movement.

As an example, suppose I am inspecting my guitar to see whether it got damaged during the process of moving my stuff to a new apartment – that seems unlikely to me, because it was carefully packaged. I just finished examining the front side of the guitar, and I am about to turn it around to take a look at its backside. According to the expectation rule, to predict the degree to which my expectations at this moment will be confirmed by the experience I am about to undergo, we have to consider the following members of  $K$ ,  $V$  and  $B$ . The relevant member of  $K$  is my knowledge that my guitar was carefully packaged (and all

---

<sup>85</sup> Yoshimi’s preferred term for the members of  $V$  is “visual images”, though he uses the term “visual sensory impressions” as well (2016, 13). It seems to me that the term “visual images” carries some unwanted connotations, so I prefer to avoid it here.

the related information I have). The relevant member of  $V$  is my sensory impression of the front side of the guitar. As for  $B$ , we should focus on my awareness that I am turning the guitar around with my hands. The expectation rule takes these inputs and yields a function whose argument is  $v$ , the visual experience I am about to undergo, and whose output is the degree to which  $v$  would fulfill my present expectations. Yoshimi calls any such function an *expectation gradient* (2016, 17–18). In the case under consideration, if I see a dent on the backside of my guitar in  $v$ , the expectation gradient would yield a low number, which means that my expectations are not very well fulfilled. In contrast, if the backside of my guitar appears intact in  $v$ , the expectation gradient would yield a number close to 1.

In my terminology, a sensory impression just is a factual content.<sup>86</sup> The expectation rule thus says that the factual contents of my present visual experience, together with my background knowledge and the kinesthetic experience I am undergoing, determine the extent to which the expectational contents of my present visual experience would be fulfilled by each of the possible visual experiences I might undergo at the following moment. As Yoshimi points out, given an expectation gradient, we can define “the set of all visual experiences above some threshold  $\tau$ , which can be taken to distinguish subsequent images that are expected (that would not surprise us) from those that would surprise us” (2016, 21).<sup>87</sup> Take the sumo case. Suppose Johannes and I just saw the referee’s gesture telling the wrestlers to start the match. Given the vast difference in background knowledge between

---

<sup>86</sup> A potential point of divergence is the following. While I take factual contents to be representational, Yoshimi doesn’t explicitly say whether the same can be said of what he calls sensory impressions. In a conversation I had with him in 2017, I was told that he would prefer to remain neutral on the issue.

<sup>87</sup> See footnote 85 for what Yoshimi means by “images” in this context.

Johannes and me, Johannes' experience and mine should be associated with different expectation gradients. Hence, though a subsequent experience of seeing a sumo wrestler charging at his opponent would not surprise Johannes, it might surprise me. It follows that such a visual experience would belong to the set of above-threshold experiences only if the set is defined by the expectation gradient associated with Johannes and not defined by the one associated with me.

Yoshimi's idea enables us to characterize the expectational base in a straightforward way. I propose to simply define the expectational base as the set that contains any possible world one would encounter in one of the above-threshold visual experiences:

(Expectational Base)

Let  $g$  be the expectation gradient determined by the background knowledge  $k$ , visual experience  $v$  and kinesthetic experience  $b$  of subject  $s$  at time  $t$ . Then for any possible world  $w$ ,  $w$  is in the expectational base of  $v$  just in case that the following is true given  $g$  in  $w$ , there is a visual experience  $v^*$  such that  $v$  is succeeded by  $v^*$  at the moment just after  $t$  and  $s$  isn't surprised by  $v^*$ .

In the scenario just described, the world in which a sumo wrestler charges at his opponent belongs to the expectational base of Johannes' experience of seeing the referee's gesture at the start of the match, but the same world doesn't belong to the expectational base of my experience had at the same time.

As suggested in the previous section, however, the expectational base alone is insufficient for determining the worlds that are expectationally relevant for the person undergoing the experience. Now that we have given a more precise definition of the

expectational base by appealing to Yoshimi's expectation gradient, we are in a better position to see why an additional theoretical construct is needed. There are two main reasons for this.

First, the worlds in the expectational base are defined to be those in which one is not surprised, and what motivates the definition is the putative link between what is not surprising and what is expected. However, the link may not be as self-evident as it might seem at first. For even though the occurrence of surprise is a good indicator of violated expectations, the lack of surprise could fail to indicate the fulfillment of expectations. This means that the expectational base fails to single out exactly those worlds that are represented by the expectational contents of one's experience.

Here's an example of a possible world that seems neither surprising nor expected. Suppose I pulled back the curtains and saw that it was snowing outside. I was not surprised. But why not? Was surprise absent because I had expected to see a possible world in which it was snowing? If we answer in the positive, we would be hard-pressed to justify our answer. I had no idea that it was going to snow today, and I drew back the curtains precisely to find out what the weather was like outside. If this is a case in which snow was expected, then it is no less a case in which many other weather phenomena, such as sunlight and rain, were expected. But it is hard to make sense of the claim that I expected all these mutually incompatible weather phenomena at once. The opposite claim may thus appear more plausible: I was not surprised when I saw snow because when I drew back the curtains, I had no expectations about the weather whatsoever.

The claim that I didn't expect any kind of weather may seem more attractive, but it is no less problematic. Imagine a case in which I pulled back the curtains and saw nothing but a tornado outside – I was extremely surprised by what I saw. However, as just mentioned,

whether or not the absence of surprise presupposes the fulfillment of expectations, the occurrence of surprise presupposes the violation of expectations. In the tornado case, I would not have been surprised if I had not expected something other than a tornado. What would be the relevant expectation? The answer is probably the expectation to see some weather phenomenon of a normal type. It is the expectation that on an ordinary day, I would see snow, rain or the like, not a tornado or a tsunami.

We are then faced with a dilemma when analyzing the expectations I had upon pulling back the curtains. In one sense, I didn't have weather-related expectations, because I would be no more surprised to see snow than to see, say, sunlight. In another sense, nonetheless, I *did* have weather-related expectations, because I expected to see an ordinary weather phenomenon instead of a rare one. How should we avoid the dilemma? A natural suggestion is to say that we are actually talking about two kinds of expectations here. To flesh out the suggestion, let's distinguish between *generic* expectations and *situational* expectations.<sup>88</sup>

On the one hand, generic expectations can be triggered by a wide variety of contexts. One might have the generic expectations that tornados are unlikely to occur, that physical entities don't vanish into thin air, that plants don't walk, that four-legged animals don't fly, and that two solid entities collide with, rather than pass through, each other upon contact. In fact, many generic expectations are nearly always operative. Take the expectation that physical entities are generally incapable of sudden disappearance. Unless one is in a very

---

<sup>88</sup> I thank Jeff Yoshimi for suggesting the terms "generic" and "situational". I considered using the terms "non-situation-specific" and "situation-specific", but they were rather clumsy.

unusual environment, it is difficult *not* to have such an expectation – no matter what physical entity I see, I will be surprised if it simply vanishes in front of my eyes.

On the other hand, situational expectations are highly context-dependent. In the sumo case, if Johannes sees one of the wrestlers grab the other's belt, he might expect a belt-throw. Such an expectation depends on the factual contents of Johannes' experience: Johannes wouldn't have the expectation if he weren't observing the sumo wrestler's action. This makes the expectation radically different from generic expectations – the generic expectation that physical objects don't suddenly vanish can be held even when one is not undergoing any visual experience. Furthermore, Johannes' expectation can only be triggered by the particular kind of situation Johannes is in. Unlike the generic expectation that plants don't walk, which is at work when one sees a chestnut tree, a maple tree, a bush or a flower, Johannes' expectation about the belt-throw is operative only when he sees a belt-grabbing action performed in a very specific way.

It is therefore sensible to distinguish between generic and situational expectations. With the distinction in hand, we can make sense of the case of pulling back the curtains. When I drew back the curtains, I had the generic expectation that I would see an ordinary weather phenomenon, such as snow or rain. What I didn't have was the situational expectation that my visual experience of the curtains would be followed by the visual experience of snow. It then emerges that only situational expectations are capable of functioning as the expectational contents of an experience – only situational expectations depend on the factual contents of an experience. Since our definition of the expectational base doesn't enable us to separate the worlds compatible with one's generic expectations

from those compatible with one's situational expectations, the expectational base fails to specify the expectational contents of one's experience on its own.

The second reason that we need additional theoretical constructs is that some worlds are included in the expectational base of one's experience *not* because they contain events that are possible continuations of the event currently perceived by one. Suppose I am watching an award ceremony. My eyeglasses are dirty, so I move my hands to take them off and wipe them clean. Given my visual experience, kinesthetic experience and background knowledge at this point, I expect to see nothing but blurry patches at the next moment. According to our definition above, a world in which only blurry patches appear is in the expectational base of my current visual experience. Such a world obviously contributes nothing to my awareness of the completing boundaries featured by the subevents of the ceremony, but the very reason we hope to identify the expectationally relevant worlds is to account for such awareness. Or consider a less extreme case: bored by the ceremony, I turn my head to talk to my friend, who sits behind me. Our definition above would say that a world in which my friend's face is located at the center of my visual field is in the expectational base of my current visual experience. Such a world tells us nothing about the way I perceive the subevents of the ceremony.

The expectational base often contains such irrelevant worlds because one of its main functions is to capture what appears possible given one's background knowledge. One's background knowledge doesn't just carry information about the kind of event one is perceiving – it also carries information about the kind of place one finds oneself in, the kind of people around one and the non-event entities currently perceived by one, etc. As a consequence, many possible worlds unrelated to the event unfolding in front of one count as

members of the expectational base. This again prevents the expectational base from independently fixing the expectational contents.

To sum up, there are three possible scenarios in which one would not be surprised by what one sees: the absence of surprise indicates the fulfillment of either (1) a generic expectation, (2) a situational expectation unrelated to the currently perceived event, or (3) a situational expectation related to the currently perceived event. In trying to identify the expectational contents of event-representing experiences, we should focus only on the possible worlds in which expectations of the third kind are fulfilled. To single out those possible worlds, we must combine the expectational base with the salience constraint. In the following section, we will turn to the latter.

#### *5.4.4 Specifying Expectational Contents: The Salience Constraint*

We have given an intuitive characterization of the way the salience constraint works, but the idea obviously needs to be further developed. The best way to proceed is to work with an example. And it would be helpful to consult some empirical work here – doing so could make our example more realistic. According to several studies, expert tennis players can visually anticipate many types of shot.<sup>89</sup> A study conducted by Jaeho Shim, Les Carlton and Young-Hoo Kwon even suggests that there could be a specific visual cue used by expert players to differentiate between lobs and groundstrokes: the pattern of relative motion featured by the opponent's forearm and racket (Shim, Carlton, and Kwon 2006, 338; also see 332 fig. 3). The pattern was measured roughly as follows (2006, 328). First, the researchers

---

<sup>89</sup> For example, see Shim et al. (2005) and Ward et al. (2002).



recorded the way a skilled player swung the racket and then broke the video into frames. Second, for each frame, they measured how much the player's forearm and racket were tilted, respectively. Third, for each frame, they calculated the ratio between these two tilt angles; on the basis of the results, they identified the pattern in which this ratio changed throughout the video. It turns out that there were considerable differences between the pattern of change in a lob and the pattern of change in a groundstroke, which suggests that these two types of stroke involve different ways of moving one's forearm relative to the racket (2006, 330).

In view of the research done by Shim and his colleagues, let's imagine the following scenario. Hannah and Edith, both experienced tennis players, were playing tennis against each other. Over time interval  $I_1$ , Edith swung her racket, and the pattern of relative motion exemplified by her forearm and racket was characteristic of a lob. Over  $I_2$ , Edith successfully hit the ball, which then moved past the net in a flight path characteristic of a lob. Let "Swing" and "Flight" denote the events that unfolded, respectively, over  $I_1$  and  $I_2$ . Call the whole event that consisted of Swing and Flight "Lob"; Lob unfolded over  $I$ , the mereological fusion of  $I_1$  and  $I_2$ . Now consider Hannah's visual experience during  $I_1$ . Hannah was skilled in tennis, so compared with novice players, she was more capable of predicting the type of stroke her opponent was going to perform on the basis her opponent's bodily movements. When Hannah saw the pattern of relative motion in Edith's swing, she expected Edith's stroke to be a lob.

I claim that Hannah's perceptual experience during  $I_1$  represented Lob as ongoing: in every world that was expectationally relevant for Hannah, the event perceived during  $I_1$  (Swing) would be succeeded by another event (Flight) that was capable of jointly making up

a complex event (Lob) with Swing.<sup>90</sup> For this reason, it appeared to Hannah that Lob had not run its course and its completing boundary was not about to occur. But why was it the case that Swing would be followed by Flight in every expectationally relevant world? That could not have been guaranteed by the expectational base of Hannah's experience during  $I_1$ . Despite Hannah's kinesthetic experience, background knowledge, and the factual contents of Hannah's perceptual experience that represented Swing, it wouldn't be too surprising for Hannah to see something other than Flight during  $I_2$ . Given how experienced Hannah was, she must have seen all kinds of happenings on the tennis court: what appeared to be a lob turned out to be a groundstroke, a swing that looked perfect ended up missing the ball, and the opponent tripped when she wasn't supposed to, etc. Arguably the worlds where these would happen to Edith all belonged to the expectational base of Hannah's experience.

The solution to this problem, as suggested above, is to invoke the salience constraint. The following notion will come in handy in our attempt to give a more precise definition of the salience constraint:

(Salient Type)

Suppose  $K$  is an event-kind and  $v$  a visual experience whose subject is  $s$  and whose object is event  $e$ . Then  $K$  is the *perceptually salient type*, or simply *salient type*, of  $e$  in  $v$  if and only if:

---

<sup>90</sup> For the reason that Lob, instead of Swing, is said to have been represented as ongoing, see footnote 80.

- (1) when undergoing  $\nu$ , it looks to  $s$  as if  $K$  is one of the event-kinds under which  $e$  falls;<sup>91</sup>
- (2) in carrying out the task at hand when undergoing  $\nu$ ,  $s$  is disposed to prioritize the information that  $e$  falls under  $K$  over the information that  $e$  falls under any other event-kind.

In the visual experience Hannah had over  $I_1$ , the salient type of Swing was the event-kind that may be characterized as *the initial phase of a lob* (henceforth abbreviated as IPL). This was probably not the only event-kind attributed to Swing by Hannah. Depending on the context, the other event-kinds Swing could appear to fall under in Hannah's experience included *ball-hitting action*, *right-arm movement* (if Hannah had just played against a left-handed player), and *human activity* (if Hannah had just practiced with a ball machine), among others. But given Hannah's task at hand, i.e. returning the ball, Hannah presumably prioritized the information that Swing fell under IPL over the information that Swing also fell under those other event-kinds. For if Hannah had had no idea how to effectively respond to Swing at all, the awareness that Swing was an instance of right-arm movement or human activity would not have been helpful. The information that Swing was a ball-hitting action was more useful, but it was not specific enough. As a consequence, among the various pieces of information about event-kinds that were carried by the factual contents of Hannah's perceptual

---

<sup>91</sup> When reporting the contents of an experience, I will sometimes follow Charles Siewert in using the locution "it looks to me as if  $\varphi$ ", where  $\varphi$  is a sentence. Siewert describes two kinds of linguistic constructions that may be used to report an experience. Suppose one is looking at the uppercase letter "X". To report this experience, one could say "It seems to me the way it does for it to look as if there's an X on my left"; alternatively, one could say "It looks to me as if there's an X on my left" (Siewert 1998, 86). Here I borrow the latter. Siewert himself prefers the former, but the subtle difference between the two need not concern us here.

experience, Hannah probably only made use of the information about IPL in returning the ball. It follows that IPL was the salient type of Swing in Hannah's experience during  $I_1$ .

With the notion of a salient type in hand, we can now define the salience constraint:

(Salience Constraint)

Let  $v$  be a visual experience of event  $e$  that subject  $s$  undergoes at time  $t$ . Say that for any possible world  $w$ ,  $w$  satisfies the salience constraint imposed by  $v$  if and only if the following are the case in  $w$ :

- (1) There is visual experience  $v^*$  such that  $v$  is succeeded by  $v^*$  at the moment just after  $t$  and  $s$  isn't surprised by  $v^*$ ;
- (2) The fact that  $s$  isn't surprised by  $v^*$  at the moment just after  $t$  obtains specifically in virtue of  $s$ 's awareness of  $e$ 's salient type in  $v$ .

Condition (1) is straightforward. But what does condition (2) mean? Let's again focus on Hannah's experience during  $I_1$ . Given that Swing was represented by the factual contents of her experience, Hannah might have expected the following to happen over  $I_2$ : the distance between the ball and the net first decreased and then increased. However, even in a world where such an expectation was indeed fulfilled during  $I_2$  (such as the actual world), the fact that Hannah wasn't surprised by her experience during  $I_2$  cannot be said to have obtained specifically in virtue of Hannah's awareness of Swing's salient type. For in a world where Swing had been replaced by an event in which Edith performed a cross-court shot, a half volley or a smash instead of a lob, Hannah would have had the same expectation. In fact, she would have had the same expectation so long as she saw Edith swinging the racket to hit the ball; it didn't matter whether the swing indicated a possible lob, a possible cross-court shot

or something else. The fact that IPL appeared to be the salient type of Swing in Hannah's experience during  $I_1$  thus contributed little to her expectation that the ball would first approach the net before flying away from it.

The above observation suggests a way to further clarify condition (2). Suppose that after undergoing visual experience  $v$ , whose object is event  $e$ , subject  $s$  fails to be surprised by the succeeding experience  $v^*$ . Let's say that the fact  $s$  isn't surprised by  $v^*$  obtains *specifically in virtue of  $s$ 's awareness of  $e$ 's salient type in  $v$*  if and only if the following holds: if  $s$  were not aware of  $e$ 's salient type in  $v$ ,  $s$  would be surprised by  $v^*$ . In the actual world, after undergoing her experience of Swing, Hannah saw the ball move towards the net and then move away from it. She wasn't surprised by this. But would Hannah have been surprised if she had not been aware of Swing's salient type? No. If Edith had performed a smash, Hannah would not have been aware of Swing's salient type, i.e. IPL. Even in that case, however, Hannah would not have been surprised to see the way the ball moved relative to the net. Hence, though there was indeed the fact that Hannah wasn't surprised by the way the ball moved relative to the net, the fact didn't obtain specifically in virtue of IPL. As another example, consider the fact that Hannah wasn't surprised by her experience of Flight. This fact obtained in the actual world, and it did obtain specifically in virtue of Hannah's awareness of IPL. If Hannah hadn't been aware of IPL, she couldn't have expected the ball to move in a path characteristic of a lob. It thus follows from our definition that the actual world satisfies the salience constraint imposed by Hannah's experience of Swing.

I shall argue in section 5.4.5 that the notions of the expectational base and the salience constraint supply the necessary tools for us to conceptualize our awareness of completing boundaries in a satisfactory manner. But first things first – I need to defend my

definition of the salience constraint. After all, one could ask: why should we focus on salient types when looking for salience constraints? An event may have many properties in addition to the properties it has in virtue of its salient type. In our tennis example, Hannah could perceive Edith's swing as performed with a red racket, in a stadium with a large audience, or carried out by a newly met opponent, etc. Swing's doesn't instantiate these properties because of its salient type – Hannah is unlikely to prioritize the information that Edith's swing is an action carried out by a newly met opponent, for example. But why not invoke these properties in BRV? Wouldn't Hannah's awareness that she is playing against a newly met opponent make a difference to the expectational contents of Hannah's experience? My answer is that, though properties like the ones just mentioned may be indeed affect one's expectations in general, they have a very limited role in determining one's expectations *of the progression of an event*. I think this claim can be vindicated on phenomenological grounds, but phenomenological disputes are rarely easy to settle. To provide additional support for my claim, therefore, I shall argue that it fits well with a sophisticated theory in empirical psychology.

The theory I have in mind is the *event segmentation theory* (EST) proposed by Jeffrey Zacks and his colleagues (Zacks et al. 2007). EST aims to identify the factors that affect how people parse a stream of sensory stimuli into multiple events. According to the theory, the mechanism responsible for event perception is predictive in nature. It mainly functions to predict what happens next given what is perceived to be happening now (Zacks et al. 2007, 273–74). For example, if I see a traffic light turning yellow now, my perceptual system might deploy the mechanism to predict that the traffic light will turn red at the following moment. How does the perceptual system make such predictions about incoming stimuli? It does so

on the basis of two kinds of representations. The first is obvious: the perceptual system extracts information from representations of current stimuli. But EST also posits a more special kind of representations – *event models*. According to the researchers, “An event model is a representation of ‘what is happening now,’ which is robust to transient variability in the sensory input” (Zacks et al. 2007, 274). The event model activated or created by one’s perceptual system in response to a set of stimuli depends in part on one’s *event schemata*, which are “semantic memory representations that capture shared features of previously encountered events” and “contain previously learned information about the sequential structure of activity” (Zacks et al. 2007, 275). Hence, in the aforementioned sumo case, Johannes’ perceptual system and mine activate very different event models in response to the perceived actions of the wrestlers, despite the fact that the sensory stimuli reaching our perceptual systems may be quite similar. Johannes’ perceptual system might predict that the wrestlers are going to run towards each other, but mine certainly makes no such prediction. This is the case because the event schemata stored in Johannes’ memory are much more detailed than those stored in my memory.

How does EST account for the way people perceptually identify the boundaries between events? According to the researchers, our perceptual system has an error-detection component. After predictions about incoming stimuli are generated at time  $t$ , they are compared with the actual stimuli received at the moment succeeding  $t$ . The perceptual system then makes adjustments to the event model in place (Zacks et al. 2007, 274–75). If the error, i.e. the discrepancy between the predicted and actual stimuli, is not significant, the perceptual system retains the current event model. By contrast, if the error is significant, the perceptual system modifies the event model in light of the actual stimuli. This might lead to

the activation of an alternative event model or the creation of one that is unlike any in the past. Significant error also results in the awareness that the transition from one event to another has occurred – this is the point at which a perceiver locates the boundary between events. Summing these up, the researchers say, “the system alternates between long periods of stability and brief periods of change. Periods of stability are perceived by observers as events, and periods of change are perceived as the boundaries between events” (Zacks et al. 2007, 275).

To make the idea more concrete, consider a connectionist network built by Jeremy Reynolds, Jeffrey Zacks and Todd Braver (2007) to test EST computationally. The stimuli used to train the network consisted of motion capture data. They depicted the various actions of a human figure, in particular the movement trajectories of 18 locations on the figure’s body. The spatial locations of these bodily locations were encoded by the units in the input layer of the network. The network had a straightforward task: given the stimuli received by its input layer at time  $t$ , the network had to predict the stimuli that would be received at the moment just after  $t$  (Reynolds, Zacks, and Braver 2007, 618). In other words, given where the figure’s bodily locations were at  $t$ , the network had to predict where they would be at the moment just after  $t$ . This was to be done with the help of a special component, an *event layer*, which was intended to be a computational mechanism realizing the event models posited by EST.



To see the role played by the event layer, consider the network's four main steps of information processing.<sup>92</sup> First, the network received initial stimuli at time  $t$ , which were represented by the activity of the input layer. Second, signals were transmitted to the event layer, processed, and sent to the output layer. The predictions made by the network were represented by the activity of the output layer at this point. Third, the network received new stimuli at the moment just after  $t$ ; these were represented by the activity of the input layer at the same time. Fourth, an error detection component was activated, which functioned to compare the new stimuli with the earlier predictions, thereby determining whether the difference between them exceeded the threshold. What was the event layer's contribution to the whole process? The answer is: its distinctive mechanism of combining signals from two sources, i.e. the input layer and *itself* (Reynolds, Zacks, and Braver 2007, 629). Notably, it was allowed to receive signals from the input layer at time  $t$  only if there was significant difference between the stimuli received at  $t$  and the predictions made at the moment just before  $t$ ; the event layer would be insulated from the input layer if the difference didn't exceed the threshold.<sup>93</sup> When prediction error was limited, therefore, the state of the event layer at a time was mainly determined by its own state at the preceding moment. According to Reynolds and his colleagues, the network was quite successful in predicting how the human figure would move its body in the next phase of the action (Reynolds, Zacks, and Braver 2007, 635).

---

<sup>92</sup> I have set aside many details that need not concern us now. See the authors' description of Simulation IV (Reynolds, Zacks, and Braver 2007, 635–36).

<sup>93</sup> More precisely, the relative weight of the signals from the input layer was adjusted in accordance with an equation whose variables included the amount of error. See (Reynolds, Zacks, and Braver 2007, 630).

What kind of information is used by networks like the one just described? I would argue that the relevant information is precisely information about salient types. Note that the activities of the event layer, which are computational realizations of the states the event model is in, are self-sustaining for most of the time. Once an event model is activated, how its state is going to change from moment to moment is largely determined by the information it carries on its own. If Edith performs the same stroke 5 times with 5 rackets of different colors, the beginnings of the 5 events are largely similar. If one is observing Edith's actions, the initial similarity between the 5 events should enable one's perceptual system to deploy the same event model in predicting what Edith is about to do – this is the case despite the individual differences between the 5 events. If the predictions so made are largely accurate, the color of Edith's racket could fail to play any role in determining one's expectations throughout the process of observing Edith's actions. It then emerges that predictions based on event models typically fail to be sensitive to the individual differences between events that largely resemble each other. What such predictions *are* sensitive to are patterns that can be repeatedly experienced in roughly the same way.<sup>94</sup> This explains why such properties as *being performed by a newly met opponent* play a negligible role in one's expectations of an event's progression – there are too many ways in which an action performed by a newly met opponent could appear to one. It could look like a cross-court

---

<sup>94</sup> Note that such patterns need not be those that have been repeatedly experienced over a long period of time; as Zacks and his colleagues point out, “Event models are working memory representations, which are implemented by transient changes in neural activation rather than long-term changes in synaptic weights” (Zacks et al. 2007, 274). Even if I am watching a tennis game for the first time in my life and I have only seen one backhand volley so far, it is possible that my perceptual system has already begun to build an event model for volleys that will affect my subsequent experiences of backhand volleys.

shot, a clumsy move or a swift action, etc. There isn't a regular pattern in which such actions are typically experienced.

What are the relevant patterns, then? It seems likely that they just are salient types. A salient type is an event-kind, and an event-kind can be plausibly understood as an abstract entity prescribing the successive stages an event must go through to be a token of the kind. The reason that, say, an event is a dive is precisely that it consists of a sequence of subevents – the subevent of standing at a diving board is followed by jumping and then falling, etc. Any other event that features the same manner of progression falls under the same event-kind as the dive. Given this, if certain patterns that are repeatedly observable can be discerned in numerically distinct events, these events can be reasonably understood to fall under the same event-kind. We can then say that event models facilitate the prediction of event progression by supplying information about event-kinds.<sup>95</sup> But not any event-kind would do. There are many kinds an event might fall under at once, and only the kind that the perceiver pays attention to is relevant. In light of this, we should say that event models do not simply carry information about some event-kind under which the currently perceived event falls – they carry information about the salient type of the currently perceived event. It then seems that EST supports the claim that salient types assume a more prominent role

---

<sup>95</sup> Some might think that the conception of event-kinds employed here is too restrictive. There may be some types of events whose tokens do not share a common pattern of progression, but there is no *a priori* reason not to count these types as event-kinds. If this is true, we could accommodate this by adopting a broader conception of event-kinds and modifying the claim about event models accordingly. Instead of simply saying that event models represent event-kinds, we say that they represent event-kinds of a particular sort, i.e. those types of events that prescribe a principle of progression governing their tokens.

than anything else in determining the expectational contents of event-representing experiences. My proposed definition of the salience constraints thus makes sense.

#### 5.4.5 *BRV Restated*

Recall our definition of expectational relevance in section 5.4.2: given a visual experience, the worlds that are expectationally relevant to the experience are those in the expectational base of the experience that meet the salience constraint imposed by the experience. We have now examined the two components of this definition, so let's give one more example to show how the definition is to be applied. Carl and I are both observing a chemical experiment, in which some red liquid is poured into a flask that holds some blue liquid. Carl is a chemist, and he has seen flasks being destroyed by liquids. The expectational base of Carl's experience thus contains a world in which the flask is destroyed at the next moment. This world does not belong to the expectational base of my experience, as I am completely ignorant in chemistry. Now, despite the fact that there are more worlds in the expectational base of Carl's experience than that of mine, the salience constraint imposed by Carl's experience is more stringent than that imposed by mine. Given the behavior of the two liquids that Carl is now observing, the only possible scenario that wouldn't surprise him is one in which certain salt-like crystals appear in the flask. It follows that the salience constraint imposed by Carl's experience excludes all the worlds in which such an event of crystallization fails to occur. Those worlds would nonetheless meet the salience constraint

imposed by my experience. As a consequence, the worlds that are expectationally relevant to my experience surpass those expectationally relevant to Carl's experience in number.<sup>96</sup>

As of yet, it remains unclear what we should say about the relation between the expectational relevance of a world and the representation of the world by expectational contents. That is, if a possible world counts as expectationally relevant, does it follow that the perceiver represents some of the occurrences in that world in the expectational contents of her experience? It is time to address this issue. Imagine a scenario in which Swing and Flight occurred the way they actually did except that Gerda, instead of Hannah, was playing against Edith. Unlike Hannah, for Gerda, it didn't seem impossible for Edith to miss the ball. This means that some worlds in which Edith missed the ball belonged to the expectational base of Gerda's experience of Swing. In those worlds, Flight failed to occur over  $I_2$ ; let "Flight-Alternatives" denote the possible events that occurred over  $I_2$  in those worlds. But neither was it obvious to Gerda that Edith would miss the ball, so the expectational base of Gerda's experience of Swing also contained worlds where Flight did occur over  $I_2$ . What, then, did the expectational contents of Gerda's experience of Swing represent? Did they represent Flight, one of the Flight-Alternatives, all of the Flight-Alternatives, a combination of Flight and Flight-Alternatives, or none of the above?

---

<sup>96</sup> Note that even if a world in the expectational base of an experience fails to be expectationally relevant, it doesn't follow that it plays no role in determining the expectational contents of the experience. It only follows that it plays no role in determining the expectational contents that *enable the perceiver to represent completing boundaries*. In the present example, though a world in which the flask is later destroyed fails to be expectationally relevant to Carl's experience, we might still have to take the world into consideration if we intend to describe Carl's perceptual experience of the flask. The world is irrelevant only with respect to how Carl perceptually represents the point at which the chemical reaction has run its course.

Here's my suggestion: the expectational contents of Gerda's experience of Swing represented an event  $e$  such that it looked to Gerda as if  $e$  could turn out to be either Flight or one of the Flight-Alternatives. In other words, the expectational contents jointly represented an event whose properties remained to be fully identified in the future. It is important not to misunderstand my claim here: I am *not* saying that there were *multiple* possible events represented by the expectational contents. Rather, there was only *one* event jointly represented by the expectational contents, and the different properties of the event were represented with different degrees of specificity. To make the point more precise, let's revisit two ideas brought up earlier. First, given an event, its *individuating property* is the conjunction of all its non-relational properties.<sup>97</sup> One component of Swing's individuating property is the property *falling under IPL*, for example. Second, *determination* is the relation between such properties as *being three-dimensional* and *being a cube*.<sup>98</sup> According to Yablo, if we focus on the metaphysical core of the determination relation, the relation can be defined as follows:

- $P$  determines  $Q$  ( $P > Q$ ) only if:
- (i) necessarily, for all  $x$ , if  $x$  has  $P$  then  $x$  has  $Q$ ; and
  - (ii) possibly, for some  $x$ ,  $x$  has  $Q$  but lacks  $P$ . (Yablo 1992, 252)

If we adopt a stronger version of Yablo's definition and turn it into a biconditional, then the definition allows us to say, for example, that the property *being a lob* determines the property

---

<sup>97</sup> See section 2.4.3.

<sup>98</sup> See section 2.4.2.

*being a tennis stroke*. After all, no action can be a lob without being a tennis stroke, but many tennis strokes are not lobs.

These two notions enable us to address the issue confronting us. Take the individuating properties of Flight and all Flight-Alternatives. Let  $E$  be the property determined by all these individuating properties. My suggestion can now be restated as follows: what was represented by the expectational contents of Gerda's experience of Swing was the event whose individuating property was  $E$ . What properties did the event have, exactly? Though we cannot fully answer this question without considering additional details, it is reasonable to assume that the property *being the movement of a tennis ball* was a component of  $E$  – since Gerda expected to see the tennis ball move no matter Edith ended up hitting the ball or not, every Flight-Alternative featured ball movement. To be sure, the property *being the movement of a tennis ball* is quite indeterminate, but a less indeterminate property is unlikely to have been a component of  $E$ . Given that the tennis ball could have moved in many directions without surprising Gerda, Flight-Alternatives involved many kinds of ball movement. If so, it is quite inconceivable that  $E$ , the property determined by the individuating properties of all Flight-Alternatives, could have comprised the property of being a particular kind of ball movement. Hence, in undergoing her experience of Swing, it looked to Gerda as if the next phase of the event would be a ball movement, but it wasn't clear to her what kind of ball movement the event would be – it wasn't clear whether the ball would move towards the net or the service line, for example. The next phase of the presently perceived event was represented by the expectational contents of Gerda's experience as a ball movement whose direction remained to be seen.

What emerges is that identifying  $E$  this way has an important implication: the more considerably Flight-Alternatives differed with each other in a given aspect, the more indeterminately that aspect was represented in the expectational contents of Gerda's experience of Swing. This observation suggests that my criterion captures an important difference between an amateur's experience of an event and an expert's experience of the same event. Recall the example in which Carl and I are observing the same chemical reaction. Since an event of crystallization occurs in every world that is expectationally relevant to Carl's experience, the expectational contents of Carl's experience represent an event of crystallization. In contrast, the event that is the object of the expectational contents of my experience is represented in a more indeterminate way. Given that my expectations of what might happen in a chemical experiment are largely based on the Hollywood movies I have watched, the expectational base of my experience contains worlds in which the chemical reaction leads to an explosion and those in which it results in leaked poisonous gas of a bizarre color. Neither type of world is excluded by the salience constraint imposed by my experience, which is much weaker than that imposed by Carl's experience. But a world in which I see an explosion is quite different from one in which I see some gas of a bizarre color. The worlds that are expectationally relevant to my experience thus conflict with each other. As a result,  $E$ , i.e. the property determined by the individuating properties of all the events that succeed the presently experienced one in the worlds expectationally relevant to my experience, is extremely indeterminate. The event whose individuating property is  $E$  would be such that something happens to the flask in the event – this much can indeed be said about that event. Except for this, however, there are probably not many determinate properties that the event would be represented to instantiate. It follows that the



expectational contents of my experience are quite impoverished: they collectively represent something that is about to happen to the flask but carry little information of what exactly that event is.

It is now time to put the various pieces of BRV together and restate it:

(BRV)

Let  $v$  be a visual experience whose factual contents represent event  $e$ .

- (1) For any event  $e^*$ ,  $e^*$  is the event represented by the expectational contents of  $v$  if and only if  $e^*$  has the individuating property  $E$ , where  $E$  is the property determined by the individuating properties of all the events that succeed  $e$  in the worlds that are expectationally relevant to  $v$ .
- (2) Let  $f$  be the event represented by the expectational contents of  $v$ . Then:
  - (a)  $v$  represents  $e$  as ongoing if and only if  $f$  jointly makes up a larger event with  $e$ . In this case,  $s$  experiences the completing boundary of  $e$  as yet to occur.
  - (b)  $v$  represents  $e$  as ending if and only if  $f$  doesn't jointly make up a larger event with  $e$ . In this case, the currently perceived temporal part of  $e$  appears as the completing boundary of  $e$ .<sup>99</sup>

This sums up the ideas we have been exploring. Take Hannah's experience of Swing as an example. Given any world  $w$  that is expectationally relevant to Hannah's experience, Flight is a part of the Flight-Alternative that occurs in  $w$ . On the assumption that the Flight-Alternatives in those worlds do not share any other part, Flight just is the event represented

---

<sup>99</sup> Note that what is said in footnote 80 still applies (the footnote can be found on p.210).

by the expectational contents of Hannah's experience. Since Hannah is well aware that Swing naturally leads to Flight, it looks to Hannah as if there is an event that Flight jointly makes up with Swing, i.e. Lob. According to BRV, Hannah perceives Lob as ongoing.

But what does it take for two events to appear capable of jointly making up a larger event with each other? My suggestion is that whether two events so appear depends on what Zacks and his colleagues call "event schemata". As noted in the previous section, event schemata belong to the long-term memory; they encode event properties including "distinctive physical features such as object and actor movement, statistical information about which patterns of activity are likely to follow a given pattern, and information about actors' goals" (Zacks et al. 2007, 275). It seems to me that the information encoded by event schemata is sufficient for distinguishing an event that can appropriately function as an additional temporal part of the presently perceived event from an event that cannot. Since event schemata are one of the factors affecting the perceptual system's selection of event models, it is likely that they are responsible for the perceiver's conscious awareness that two events are, or are not, temporal parts of the same complex event. We can now give another example to illustrate BRV. Suppose that Carl has conducted the aforementioned chemical experiment countless times in his lab and his lab has a machine that automatically empties the flask immediately after crystallization is detected. Let's focus on Carl's visual experience of the crystallization event. Together with the factual contents of this experience, Carl's kinesthetic experience, background knowledge and awareness of the chemical reaction's salient type yield a set of expectationally relevant worlds whose members all contain an event in which the flask appears emptied. According to Carl's event schemata of both the event of crystallization and the emptying of the flask, the latter doesn't jointly make up a single event

with the former. BRV thus entails that Carl perceives the chemical reaction as ending with the crystallization event, which appears to be the completing boundary of the chemical reaction.

An important point to note here is that though I have been working with examples that involve specialized skills, BRV is by no means a theory that aims exclusively to account for specialized skills. I am unfamiliar with tennis or chemical experiments, but I am familiar with riding a bike, climbing a ladder, drinking beer or turning the pages of a book. When I see people engaging in these everyday activities, the expectational contents of my experience can be quite detailed. The awareness of completing boundaries is a pervasive phenomenon when it comes to the events people perceive in their everyday life, and this point is properly acknowledged by BRV.

One might be worried that BRV doesn't really mesh with EST. According to EST, a boundary between events is perceived when the predictions made by the perceptual system depart considerably from what is subsequently seen. In other words, the awareness of boundaries results from the frustration of expectations. But BRV says that one can experience completing boundaries even *before* such frustration occurs. Doesn't this contradict EST? Here's my response. While EST can be plausibly viewed as a theory of what it takes for one to see an event as *having ended*, BRV is better considered as a theory of what it takes for one to see an event as *ending* or *about to end*. It is one thing to see a move as the final move in a ballet dance and another to see that a ballet dance has ended and a new dance has begun. BRV is mainly a theory of the former, and by having such a focus, it captures certain differences between experts' and amateurs' experiences that EST pays less attention to. As the cases of Hannah and Carl indicate, within the framework of BRV, one experiences an

event as ongoing or ending only if the expectational contents of one's experience are quite rich. While experts like Hannah and Carl may experience a tennis stroke or a chemical reaction as ongoing or ending, it is less likely for ordinary people to undergo such perceptual experiences. I have suggested that if I were to undergo a perceptual experience of the chemical experiment, the event that would be represented by the expectational contents of my experience would putatively appear in a very indeterminate way – it would probably appear as something that is about to happen to the flask but nothing more determinate than that. Moreover, my event schemata of chemical reactions don't provide much useful information about the experiment I am observing. It is therefore unclear to me whether the events represented by the factual and expectational contents of my experience would jointly make up a single event. The right thing to say in such a case, I submit, is that I neither experience the presently perceived event as ongoing nor experience it as ending – my perceptual experience fails to indicate one way or the other. But we shouldn't jump to the conclusion that I cannot perceive any completing boundaries in the chemical experiment. Though I am incapable of perceiving the chemical reaction as ending, it is certainly possible for me to perceive it as having ended. The only difference is that my awareness of the completing boundary would occur later than Carl's awareness of the same completing boundary. BRV can accommodate my awareness by saying that it occurs either because what is perceived to be happening defies my earlier expectations or because the event unfolding now appears incapable of making up a single event with the event I perceived at the preceding moment. In other words, my awareness of the completing boundary occurs either because of frustrated expectations or because of perceived conflicts. Such an analysis is compatible with EST, which associates the representation of a completing boundary with the

perceptual system's reaction to significant prediction error. Though the concerns of BRV do not entirely overlap those of EST, EST complements BRV rather than contradicts it.

## 5.5 Revisiting the Central Thought Experiment

### 5.5.1 *Are All Expectational Contents Non-Perceptual Contents?*

We are now ready to revisit the thought experiment about the Newton-style study described at the beginning of the chapter, i.e. the Central Thought Experiment. Given the difference between Kasimir's and Alexius' background knowledge, Kasimir identified more boundaries between events than Alexius did even though they were watching the same video. We posed the question: did Kasimir and Alexius have different *perceptual experiences*? It is clear now how I would answer the question. Kasimir's and Alexius' perceptual experiences were indeed different, and that was the case because the expectational contents of their perceptual experiences differed. The worlds expectationally relevant to Kasimir's experience were different from those expectationally relevant to Alexius' experience. Of course, my response makes sense only if one is willing to accept a central component of HDC: the claim that perception has expectational contents. Some might be unhappy with this claim. To defend my response to the Central Thought Experiment, I shall argue for the claim about expectational contents in the remainder of the chapter.

How might one object to the claim? I take it to be uncontroversial that people familiar with an event often have expectations that are different from those had by people unfamiliar with the event. What is probably more controversial is that such expectations are sometimes contents of perception. To clarify the issue, it is helpful to invoke the distinction

between what is *constitutive* of an experience and what *causes* the experience (Hopp 2011, 53; Alston 1999, 185). According to William Alston, constitutive considerations are those about “the *intrinsic character* of experience, not about the causal influences that are responsible for that” (1999, 185). An entity constitutive of an experience is a component of the experience itself, but the causes of an experience are typically entities distinct from the experience. It is possible to frame an objection to the idea that perception has expectational contents in terms of the distinction: though expectations sometimes cause perceptual experiences that have certain contents, they are never constitutive of those perceptual contents.

I don't think the objection stands to scrutiny. To respond to the objection, let's revisit the recurring theme of tennis and consider the following argument from skills. How a professional tennis player would respond to a move by the opponent is obviously different from how a novice player would respond to the same move. How should we explain such a difference? Those rejecting expectational contents could offer the following explanation: the skilled player has the same visual experience as the novice player, but the former's dispositions to act are different from those of the latter. As a result, the two players carry out different actions in response to the same move. Such an explanation seems untenable to me. The novice player has little idea of what to look out for when the opponent is attempting to hit the ball. Given her inability to effectively direct her attention, the way the opponent swings the racket in performing a lob could look exactly the same as the way the opponent swings the racket in performing a down-the-line shot. It follows that the novice player has the same visual experience in these different scenarios. If the professional player has the same experience as the novice player, then the former also undergoes the same experience in the two scenarios. But the professional player's dispositions to act would then be insufficient

to explain why her action is different from the novice player's. Dispositions can only explain why the professional player would act in one way when seeing a move that looks like a lob but act in another way when seeing a move that looks like a down-the-line shot. If the player were to react in different ways to two moves that look exactly the same, then the player would just be acting arbitrarily. That would then entail that the player isn't disposed to act in any particular way – dispositions have no role in the explanation of the player's behavior. The appeal to dispositions thus fails to tell us why the professional player and the novice player would react differently to the same move by the opponent.

As a rejoinder, those rejecting expectational contents might say the following. Even if there is indeed a difference between the two players' visual experiences, it doesn't follow that the difference should be explained in terms of a difference in expectations. Perhaps the explanation should be that the two players attend to different aspects of the opponent's action. The first thing to note in responding to this argument is that the argument rests on more presuppositions than the previous one. While it is quite conceivable that one's expectations may remain unchanged even when one's dispositions to act have changed, it is less clear that variations in attention can occur without variations in expectations. To deny that there are expectations constitutive of perceptual contents, therefore, one has to show that the independent variation of attention is possible. I doubt that it is possible. Suppose the amateur player is told what the skilled player would focus on and asked to attentively track the aspects of the opponent's move that the skilled player would attentively track. Would the amateur player succeed? My prediction is that the player would fail. The only way for one to deploy one's attention the way a professional player does is to become a professional player; the ability to effectively deploy one's attention cannot be learned

independently of the other tennis skills. But why? My suggestion is that one has to learn to entertain the right expectations, and one's expectations depend on one's kinesthetic experiences and event schemata. Before one has mastered the relevant range of bodily skills, one is simply incapable of undergoing the necessary kinesthetic experiences and acting in accordance with the relevant event models.

To see this, imagine a counterfactual scenario similar to the one just described, except that the opponent made some additional gestures intended to confuse. The professional player was able to ignore the gestures: she expected the gestures to contribute nothing to a successful shot. In light of such expectations, the professional player was able to direct her attention elsewhere upon seeing the gestures. But the amateur was confused by the gestures. Given that she had no idea what to expect given what she was seeing, she had to pay attention to any detail she happened to see, whether the detail could be safely ignored or not. As a consequence, her use of attentional resources was much less effective than that of the professional player's. This suggests that a genuine ability to deploy one's attention is premised upon one's ability to entertain the relevant kind of expectation. If so, the reason that professional and novice players direct their attention in different ways when seeing the same move may well be that different expectations are elicited when they see the move. We should therefore uphold the thesis that perception has expectational contents. The fact that tennis players with different levels of expertise attend to the same move in different ways is likely to support the thesis rather than contradict it.

But perhaps there is yet another argument that those rejecting expectational contents could make. They could argue that though the same event is represented differently by the



two tennis players in our example, the representational difference can be explained in terms of beliefs. In the next section, I discuss the reasons that such a view is objectionable.

### 5.5.2 *Are All Expectational Contents Belief Contents?*

I take the issue at hand to be more circumscribed than the heatedly debated issue of the conceptual contents of perception. Philosophers have conflicting views on the extent to which perceptual contents are conceptual, and no consensus has been reached on what kinds of contents count as conceptual. Here I am only interested in one variety of conceptual contents: the contents of beliefs. I will not take a stand on whether all belief contents are conceptual or whether all conceptual contents are belief contents. My only concern is to argue against the idea that all expectational contents are belief contents. Given that the topic here is not the nature of beliefs, I will simply assume that there is an intuitive distinction between perceptual experiences and beliefs. As A. D. Smith observes in discussing the Müller-Lyer illusion, “Being familiar with the Müller-Lyer illusion, I do *not* believe its two principal lines to be unequal, though they certainly do look it” (2001, 287). In this case, though one’s visual experience of the Müller-Lyer diagram is accompanied by a belief, the contents of the experience are different from the contents of the belief. One’s belief and experience represent the same object to exemplify different properties.

How might one go about arguing that all expectational contents are belief contents? An intuitively appealing argument can be found by modifying what Hopp calls “the argument from perceiving-as”, whose key premise can be stated in one sentence: “if every object of perception is conceptualized, then perception must always have conceptual content” (2011, 47). When one sees, say, a cantaloupe, one typically perceives it as a

cantaloupe – one categorizes it as a cantaloupe. But to categorize an object is to conceptualize it. According to the key premise of the argument from perceiving-as, one's perceptual experiences of cantaloupes have conceptual contents. The same reasoning applies to any other perceptual experience. To adapt this argument to the case of expectations, one could point out that the object of an expectation always appears as being a certain way – as a lob, as a dive, as involving many people or as taking place in a tennis court, for example. If so, what one expects is always represented as being  $p$ , where  $p$  is some property. But one cannot represent an entity as being  $p$  without entertaining some beliefs about what it takes for something to be  $p$ . It follows that expectations cannot occur without co-occurring beliefs. Since nothing is lost if we go further and identify expectations with beliefs, we should make this move to achieve theoretical parsimony. We can then conclude that the contents of expectations are exhausted by belief contents.

The premise I find the most contestable in this argument is that one has to entertain beliefs about a property in order to represent something as having that property. To be sure, such a premise seems compelling: is it really possible for one to expect there to be a lob or a dive without entertaining some beliefs about what a lob or a dive is? Yes, it is. One can be said to expect a lob insofar as one expects an entity that looks like a tennis ball to move along a path characteristic of a lob. Unless we make the contentious assumption that all representations of motion, shape and color presuppose beliefs about motion, shape and color, it is possible for such an expectation to occur independently of any beliefs. And there is no reason to accept the contentious assumption. Based on the data accumulated in empirical research, the psychologist Jeremy Wolfe has compiled a list of properties that are variously described as “stimulus attributes that guide attention”, “preattentive dimensions”

and “pop-out features” (Wolfe 2014, 23). When one looks for something in one’s environment, these are the properties that can be used to make one’s searching process more efficient (2014, 20). For example, if one is asked to stare at a screen and identify the number of rotating color patches that are triangular in shape, one might use the property of rotation to differentiate those stimuli that one should pay attention to from those that one shouldn’t, i.e. the stationary color patches. If this happens, rotation is a pop-out feature, and attention isn’t necessary for the representation of pop-out features. In the list compiled by Wolfe, color and motion are among the “undoubted” pop-out features, and shape is among the “probable and possible” ones (2014, 23–25). Now, if pop-out features can be perceived without attention, it seems unlikely that the perception of such features requires beliefs about them. There are always many properties that are perceptually represented without being attended to, and one is usually incapable of entertaining beliefs about all these properties. But if one can perceive pop-out features without entertaining beliefs about them, why cannot one expect them without entertaining beliefs about them? And if beliefs aren’t necessary for the expectations of such properties, nothing prevents us from treating them as perceptual contents to accommodate the observation that acquiring such expectations makes a lot of difference to one’s visual experiences.

If the above reasoning is plausible, there are expectational contents that are not belief contents. This point is further supported by the observations offered by scholars working in the phenomenological tradition. Developing an idea found in Dahlstrom (2006) and Yoshimi (2009), Hopp argues that “Horizons... are radically situation-dependent, so much so that any two experiences with different intuitive contents must have different

horizontal contents. But concepts are not individuated that finely” (Hopp 2010, 20).<sup>100</sup>

Insofar as belief contents are conceptual contents, it is possible for me to entertain the belief “this is a watermelon” both when I see a watermelon sliced in half and when I see an intact watermelon. But it is impossible for me to have exactly the same expectations on these two occasions. This means that the kind of expectations entertained on these occasions cannot be construed as beliefs.

I think we can reinforce Hopp’s point by taking into consideration the relation between perspectives and fulfillment. A belief imposes little perspective-based constraints on the range of experiences that could fulfill the belief. This is the case not only for ordinary beliefs but also for those beliefs that consist of what Hopp calls *authentically possessed concepts*: “A person possesses a concept authentically just in case she has sufficiently reliable, nondeferential capacities to identify its object over a sufficiently wide range of conditions and environments” (Hopp 2011, 195). Take Husserl’s example of seeing a flying black bird in Investigation VI of *Logical Investigations* (Husserl [1913b] 1970, 2:195). If I possess the concepts “flying”, “black” and “bird” authentically, then my belief that the black bird I see is flying will tend to be veridical whenever it is entertained, no matter how different the entities represented by my belief may look on these different occasions. Even if I see a black raven on one occasion and a black cowbird on another, I will not be confused by their different appearances – I will not entertain a belief that represents a bird in the first case and a belief that represents a squirrel in the second case, for example. However, even such beliefs can be

---

<sup>100</sup> As mentioned in section 4.3.2, “intuitive contents” and “horizontal contents” are Hopp’s terms for what I call “factual contents” and “expectational contents”.

fulfilled by the factual contents of a wide variety of perceptual experiences that are had from different perspectives. I can look at a flying black bird that is nearby or far away, from above or below, with my right eye closed or with my sunglasses on, etc., and my belief will be fulfilled by the factual contents of all these perceptual experiences. However authentic my possession of the relevant concepts is, any experience in which I perceive a flying black bird would verify my belief *to some degree* – this is the case even though different experiences of such a bird might provide different degrees of fulfillment.<sup>101</sup> By contrast, expectations impose considerable perspective-based constraints on what perceptual experiences could fulfill them. If I see a flying black bird nearby and expect it to move towards me, there is no way an experience in which the same bird is perceived to be moving towards me from far away could possibly fulfill my expectation. Similarly, if I expect the bird to land on a tree when I undergo an experience in which I see the bird from below, it is impossible for a later experience in which I see the bird from above to fulfill my expectation, even if the later experience does represent the bird as landing on the tree. The intimate relation between perspectives and the fulfillment of expectations is simply not found when it comes to the fulfillment of beliefs.

Some might worry that there are boundary cases. For example, Madary notes that the distinction between beliefs and expectational contents isn't always clear-cut. He discusses the case in which one is told that a surprise is waiting in the room when one is about to enter the room. He argues that in this case, one can be said to expect the surprise, but one's

---

<sup>101</sup> See Hopp (2011, 121) for a discussion of how the way an object is perceived might affect the degree of fulfillment.

expectation “is like a belief in the sense that it is not, or is only minimally, constrained by one’s current perceptual context” (2016, 53). After all, one has no idea what is going to happen in the room. However, Madary’s example doesn’t show that a distinction between expectations and beliefs cannot be drawn on the basis of possible ways to undergo fulfillment. In the example, given one’s background knowledge, one would presumably expect to see events of the kinds that are typically featured by surprises: a birthday party or a special family gathering, for example. It’s just that one doesn’t expect any of these events in particular. There is nonetheless a way events of these kinds are usually experienced – one usually attends them as a participant, not, say, as a special agent sneaking into the house to complete a stealth mission. It is therefore part of one’s background knowledge that the people one is about to see will be seen from a position where one can talk to them and maintain eye contact, not from a position above people’s heads. In light of this observation, let’s add some more details to Madary’s example and suppose there are two ways to enter the room: from the door or from the chimney. As our observation suggests, there is a kind of expectation that can be fulfilled however one enters the room, but there is another kind of expectation whose fulfillment depends upon the exact way one enters the room. One’s expectation in the first case is a belief. If one entertains the belief that there are people in the room, such a belief can be fulfilled insofar as people are seen in the room – it doesn’t matter whether one enters the room from the door as a participant or from the chimney as a special agent. But one’s expectation in the second case is not a belief. If one expects to see people in the room after undergoing a visual experience of a closed door, such an expectation can only be fulfilled by a later experience in which one sees people when standing at the door. In a scenario in which one sees people’s heads from above after seeing a closed door, one would

begin to doubt one's senses. Hence, even in a case like the one described by Madary, we can still differentiate between belief contents and the expectational contents of perception. It won't do to assimilate the latter to the former.

## 5.6 Conclusion

My response to the Central Thought Experiment completes my presentation of BRV. According to BRV, the expectationally relevant worlds are those containing an event that the presently perceived event naturally transitions to in accordance with its salient type. On the basis of such worlds, we can identify the event represented by the expectational contents of one's experience. If that event is capable of jointly making up a single complex event with the presently perceived event, then the presently perceived event is experienced as ongoing. If not, then it is experienced as ending.

Several issues pertinent to my view have not been addressed, but this chapter has already been too long. They will therefore have to be addressed elsewhere. Before ending this chapter, I would like to mention one such issue. I have been deliberately using the ambiguous term "jointly making up with". How exactly should the term be understood? My preferred view is that two events jointly make up another if the former jointly constitute the latter. Nevertheless, the boundary representation view is supposed to be a phenomenological theory of event perception. Given that my view on event constitution, i.e. dependence-based hylomorphism, analyzes constitution in terms of the satisfaction of immediate and specific dependence needs, it is not easy to tell in what way the constitution relation figures in perception. I believe this problem can be solved by appealing to a suitable notion of isomorphism. We could argue that when two events appear to constitute another in

conformity with one's event schemata, there is a relation-preserving mapping from the components of the constituted event to the components of one's experience. But many details remain to be fleshed out, and I cannot explore this idea any further here.



## BIBLIOGRAPHY

- Alston, William P. 1999. "Back to the Theory of Appearing." *Philosophical Perspectives* 13: 181–203.
- Austin, J. L. (1957) 1979. "A Plea for Excuses." In *Philosophical Papers*, edited by J. O. Urmson and G. J. Warnock, 3rd ed., 175–204. Oxford: Oxford University Press.
- Baker, Lynne Rudder. 2007. *The Metaphysics of Everyday Life: An Essay in Practical Realism*. Cambridge: Cambridge University Press.
- Banick, Kyle. 2017. "Intensionality and Intentionality: Phenomenology, Logic, and Mind." PhD diss., University of California, Irvine. Retrieved from <<https://escholarship.org/uc/item/25m22937>>.
- Bennett, Jonathan. 1988. *Events and Their Names*. Indianapolis, IN: Hackett.
- . 1996. "What Events Are." In *Events*, edited by Roberto Casati and Achille C. Varzi, 137–51. Aldershot, England: Dartmouth Publishing.
- Bennett, Karen. 2017. *Making Things Up*. Oxford: Oxford University Press.
- Bennett, Michael, and Barbara H. Partee. (1978) 2004. "Toward the Logic of Tense and Aspect in English." In *Compositionality in Formal Semantics: Selected Papers by Barbara H. Partee*, by Barbara H. Partee, 59–109. Oxford: Blackwell.
- Brewer, Bill. 2011. *Perception and Its Objects*. Oxford: Oxford University Press.
- Burge, Tyler. 1977. "A Theory of Aggregates." *Noûs* 11 (2): 97–117.
- . 2010. *Origins of Objectivity*. Oxford: Oxford University Press.
- Byrne, Alex. 2001. "Intentionalism Defended." *The Philosophical Review* 110 (2): 199–240.
- Campbell, John. 2002. *Reference and Consciousness*. Oxford: Oxford University Press.
- . 2014. "Experiencing Objects as Mind-Independent." In *Berkeley's Puzzle: What Does Experience Teach Us?*, by John Campbell and Quassim Cassam, 50–74. Oxford: Oxford University Press.
- Cartwright, Helen Morris. 1970. "Quantities." *The Philosophical Review* 79 (1): 25–42.
- Casati, Roberto. 2015. "Object Perception." In *The Oxford Handbook of Philosophy of Perception*, edited by Mohan Matthen, 393–404. Oxford: Oxford University Press.

- Casati, Roberto, and Achille C. Varzi. 1999. *Parts and Places: The Structures of Spatial Representation*. Cambridge, MA: The MIT Press.
- . 2008. “Event Concepts.” In *Understanding Events: From Perception to Action*, edited by Thomas F. Shipley and Jeffrey M. Zacks, 31–54. Oxford: Oxford University Press.
- . 2015. “Events.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2015. Stanford, CA: Metaphysics Research Lab, Stanford University. Retrieved from <<https://plato.stanford.edu/archives/win2015/entries/events/>>.
- Champollion, Lucas. 2011. “Quantification and Negation in Event Semantics.” *The Baltic International Yearbook of Cognition, Logic and Communication* 6 (1): 1–23.
- . 2017. *Parts of a Whole: Distributivity as a Bridge between Aspect and Measurement*. Oxford: Oxford University Press.
- Chisholm, Roderick. 1970. “Events and Propositions.” *Noûs* 4 (1): 15–24.
- Cleland, Carol. 1991. “On the Individuation of Events.” *Synthese* 86 (2): 229–54.
- Correia, Fabrice. 2004. “Husserl on Foundation.” *Dialectica* 58 (3): 349–67.
- . 2007. “(Finean) Essence and (Priorean) Modality.” *Dialectica* 61 (1): 63–84.
- Crane, Tim. 2006. “Is There a Perceptual Relation?” In *Perceptual Experience*, edited by Tamar Szabó Gendler and John Hawthorne. Oxford: Clarendon Press.
- . 2011. “Is Perception a Propositional Attitude?” In *The Admissible Contents of Experience*, edited by Katherine Hawley and Fiona Macpherson, 83–100. Malden, MA: Wiley-Blackwell.
- Crowther, Thomas. 2011. “The Matter of Events.” *Review of Metaphysics* 65 (1): 3–39.
- . 2014. “The Perception of Activity.” *Ratio* 27 (4): 439–61.
- Dahlstrom, Daniel O. 2006. “Lost Horizons: An Appreciative Critique of Enactive Externalism.” In *Passive Synthesis and Life-World*, edited by Alfredo Ferrarin, 211–31. Pisa: Edizioni ETS.
- Dainton, Barry. 2018. “Temporal Consciousness.” In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2018. Stanford, CA: Metaphysics Research Lab, Stanford University. Retrieved from <<https://plato.stanford.edu/archives/win2018/entries/consciousness-temporal/>>.
- Davidson, Donald. (1969) 2001. “The Individuation of Events.” In *Essays on Actions and Events*, 163–80. Oxford: Oxford University Press.

- . (1967) 2001. “The Logical Form of Action Sentences.” In *Essays on Actions and Events*, 105–21. Oxford: Oxford University Press.
- Davies, Martin. 1997. “Externalism and Experience.” In *The Nature of Consciousness*, edited by Ned Block, Owen Flanagan, and Güven Guzeldere, 309–27. Cambridge, MA: The MIT Press.
- Di Nasso, Mauro. 2002. “An Axiomatic Presentation of the Nonstandard Methods in Mathematics.” *The Journal of Symbolic Logic* 67 (1): 315–25.
- Dowty, David R. 1977. “Toward a Semantic Analysis of Verb Aspect and the English ‘Imperfective’ Progressive.” *Linguistics and Philosophy* 1 (1): 45–77.
- . 1979. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague’s PTQ*. Boston, MA: Kluwer.
- Dowty, David R., Robert E. Wall, and Stanley Peters. 1981. *Introduction to Montague Semantics*. Boston: Kluwer.
- Dretske, Fred I. 1981. *Knowledge and the Flow of Information*. Cambridge, MA: The MIT Press.
- . 1995. *Naturalizing the Mind*. Cambridge, MA: The MIT Press.
- Egan, Andy. 2010. “Projectivism without Error.” In *Perceiving the World*, edited by Bence Nanay, 68–96. Oxford: Oxford University Press.
- Evans, Gareth. 1982. *The Varieties of Reference*. Edited by John McDowell. Oxford: Oxford University Press.
- Evnine, Simon J. 2011. “Constitution and Composition: Three Approaches to Their Relation.” *ProtoSociology* 27: 212–35.
- . 2016. *Making Objects and Events: A Hylomorphic Theory of Artifacts, Actions, and Organisms*. Oxford: Oxford University Press.
- Fine, Kit. 1982. “Acts, Events and Things.” In *Language and Ontology: Proceedings of the 6th International Wittgenstein Symposium*, edited by Werner Leinfellner, Eric Kraemer, and Jeffrey Schank, 97–105. Vienna: Hölder-Pichler-Tempsky.
- . 1995. “Part-Whole.” In *The Cambridge Companion to Husserl*, edited by Barry Smith and David Woodruff Smith. Cambridge: Cambridge University Press.
- Fodor, Jerry A. 2007. “The Revenge of the Given.” In *Contemporary Debates in Philosophy of Mind*, edited by Brian P. McLaughlin and Jonathan Cohen, 105–16. Malden, MA: Blackwell.

- Freeman, Jeremy, and Eero P. Simoncelli. 2011. "Metamers of the Ventral Stream." *Nature Neuroscience* 14: 1195–1201.
- Geach, Peter. 1969. *God and the Soul*. London: Routledge and Kegan Paul.
- Gramann, K., H. J. Müller, B. Schönebeck, and G. Debus. 2006. "The Neural Basis of Ego- and Allocentric Reference Frames in Spatial Navigation, Evidence from Spatio-Temporal Coupled Current Density Reconstruction." *Brain Research* 1118 (1): 116–29.
- Hintikka, Jaakko. 1969. *Models for Modalities: Selected Essays*. Dordrecht, Holland: D. Reidel.
- Hopp, Walter. 2008. "Husserl on Sensation, Perception, and Interpretation." *Canadian Journal of Philosophy* 38 (2): 219–46.
- . 2010. "How to Think about Nonconceptual Content." *The New Yearbook for Phenomenology and Phenomenological Philosophy* 10: 1–24.
- . 2011. *Perception and Knowledge: A Phenomenological Account*. Cambridge: Cambridge University Press.
- . 2016. "Empty Intentions and Phenomenological Character: A Defense of Inclusivism." In *Phenomenology of Thinking: Philosophical Investigations Into the Character of Cognitive Experiences*, edited by Thiemo Breyer and Christopher Gutland, 44–61. New York: Routledge.
- Husserl, Edmund. (1913a) 1970. *Logical Investigations*. Translated by J. N. Findlay. Vol. 1. 2 vols. New York: Routledge.
- . (1913b) 1970. *Logical Investigations*. Translated by J. N. Findlay. Vol. 2. 2 vols. New York: Routledge.
- . (1973) 1997. *Thing and Space: Lectures of 1907*. Translated by Richard Rojcewicz. Boston: Kluwer.
- . (1966) 2001. *Analyses Concerning Passive and Active Synthesis: Lectures on Transcendental Logic*. Translated by Anthony J. Steinbock. New York: Springer.
- . 2008. *Introduction to Logic and Theory of Knowledge: Lectures 1906/07*. Translated by Claire Ortiz Hill. Dordrecht, The Netherlands: Springer.
- Inwagen, Peter van. 1990. *Material Beings*. Ithaca, NY: Cornell University Press.
- James, William. 1890. *The Principles of Psychology*. Vol. 1. 2 vols. New York: H. Holt and Company.
- Johnson, W. E. 1921. *Logic, Part 1*. Cambridge: Cambridge University Press.

- Johnston, Mark. 2006. "Hylomorphism." *The Journal of Philosophy* 103 (12): 652–98.
- Jones, Tessa. 2013. "The Constitution of Events." *The Monist* 96 (1): 73–86.
- Kearns, Kate. 2011. *Semantics*. 2nd ed. New York: Palgrave Macmillan.
- Kim, Jaegwon. 1976. "Events as Property Exemplifications." In *Action Theory*, edited by Myles Brand and Douglas Walton, 159–77. Dordrecht, South Holland: D. Reidel.
- Kosslyn, Stephen M. 1984. "Mental Representation." In *Tutorials in Learning and Memory: Essays in Honor of Gordon Bower*, edited by John R. Anderson and Stephen M. Kosslyn, 91–117. New York: W. H. Freeman and Company.
- . 1994. *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, MA: The MIT Press.
- Kosslyn, Stephen M., William L. Thompson, and Giorgio Ganis. 2006. *The Case for Mental Imagery*. Oxford: Oxford University Press.
- Kratzer, Angelika. 1998. "More Structural Analogies between Pronouns and Tenses." In *Proceedings from Semantics and Linguistic Theory*, 8:92–110.
- . (1981) 2002. "The Notional Category of Modality." In *Formal Semantics: The Essential Readings*, edited by Paul Portner and Barbara Partee, 289–323. Oxford: Blackwell.
- Lewis, David. 1986a. "A Subjectivist's Guide to Objective Chance." In *Philosophical Papers Vol. II*, 83–113. Oxford: Oxford University Press.
- . 1986b. "Causation." In *Philosophical Papers Vol. II*, 159–213. Oxford: Oxford University Press.
- . 1986c. "Events." In *Philosophical Papers Vol. II*, 241–69. Oxford: Oxford University Press.
- . 1986d. *On the Plurality of Worlds*. Oxford: Blackwell.
- Liu, Feng-hsi. 2014. "Quantification and the Count-Mass Distinction in Mandarin Chinese." In *Peaches and Plums: Essays on Language and Linguistics in Honor of Rudolph C. Troike*, edited by C.-T. James Huang and Feng-hsi Liu, 153–80. Taipei, Taiwan: Institute of Linguistics, Academia Sinica.
- Madary, Michael. 2016. *Visual Phenomenology*. Cambridge, MA: The MIT Press.
- Martin, M. G. F. 2004. "The Limits of Self-Awareness." *Philosophical Studies* 120 (1/3): 37–89.
- Matthen, Mohan. 2014. "Image Content." In *Does Perception Have Content?*, edited by Berit Brogaard, 265–90. Oxford: Oxford University Press.

- McDowell, John. 1996. *Mind and World*. Cambridge, MA: Harvard University Press.
- Miller, Alexander. 2007. *Philosophy of Language*. 2nd ed. Montreal and Kingston: McGill-Queen's University Press.
- Miller, Louisa, Hannah C. Agnew, and Karin S. Pilz. 2018. "Behavioural Evidence for Distinct Mechanisms Related to Global and Biological Motion Perception." *Vision Research* 142: 58–64.
- Moltmann, Friederike. 2003. "Events as Derived Objects." In *Empirical Issues in Formal Syntax and Semantics 4*, edited by Claire Beyssade, Olivier Bonami, Patricia Cabredo Hofherr, and Francis Corblin, 189–201. Presses de l'Université Paris-Sorbonne.
- Mulligan, Kevin, and Fabrice Correia. 2017. "Facts." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2017. Stanford, CA: Metaphysics Research Lab, Stanford University. Retrieved from <<https://plato.stanford.edu/archives/win2017/entries/facts/>>.
- Needham, Paul. 2013. "Process and Change: From a Thermodynamic Perspective." *British Journal for the Philosophy of Science* 64 (2): 395–422.
- Newtson, Darren. 1973. "Attribution and the Unit of Perception of Ongoing Behavior." *Journal of Personality and Social Psychology* 28 (1): 28.
- Nolan, Daniel. 2013. "Why Historians (and Everyone Else) Should Care about Counterfactuals." *Philosophical Studies* 163 (2): 317–35.
- O'Callaghan, Casey. 2008. "Object Perception: Vision and Audition." *Philosophy Compass* 3 (4): 803–29.
- Ögmen, Haluk, and Bruno G. Breitmeyer, eds. 2006. *The First Half Second: The Microgenesis and Temporal Dynamics of Unconscious and Conscious Visual Processes*. Cambridge, MA: The MIT Press.
- Olson, Eric T. 2006. "Temporal Parts and Timeless Parthood." *Noûs* 40 (4): 738–752.
- Palmer, Stephen, and Irvin Rock. 1994. "Rethinking Perceptual Organization: The Role of Uniform Connectedness." *Psychonomic Bulletin & Review* 1 (1): 29–55.
- Pautz, Adam. 2009. "A Simple View of Consciousness." In *The Waning of Materialism: New Essays*, edited by Robert C. Koons and George Bealer, 25–66. Oxford: Oxford University Press.
- Peacocke, Christopher. 1999. *Being Known*. Oxford: Oxford University Press.

- Phillips, Ian. 2008. "Perceiving Temporal Properties." *European Journal of Philosophy* 18 (2): 176–202.
- Pollock, John. 1984. *The Foundations of Philosophical Semantics*. Princeton: Princeton University Press.
- Portner, Paul. 1998. "The Progressive in Modal Semantics." *Language* 74 (4): 760–87.
- Pylyshyn, Zenon W. 2007. *Things and Places: How the Mind Connects with the World*. Cambridge, MA: The MIT Press.
- Quine, Willard V. 1985. "Events and Reification." In *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, edited by Ernest Lepore and Brian P. McLaughlin, 162–71. Oxford: Blackwell.
- Reynolds, Jeremy R., Jeffrey M. Zacks, and Todd S. Braver. 2007. "A Computational Model of Event Segmentation from Perceptual Prediction." *Cognitive Science* 31 (4): 613–43.
- Rosefeldt, Tobias. 2018. "Should Metaphysics Care about Linguistics?" *Journal for General Philosophy of Science* 49 (2): 161–78.
- Schaffer, Jonathan. 2009. "On What Grounds What." In *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David J. Chalmers, David Manley, and Ryan Wasserman, 347–83. Oxford: Oxford University Press.
- Sharon, Tanya, and Karen Wynn. 1998. "Individuation of Actions from Continuous Motion." *Psychological Science* 9 (5): 357–62.
- Shim, Jaeho, Les G. Carlton, John W. Chow, and Woen-Sik Chae. 2005. "The Use of Anticipatory Visual Cues by Highly Skilled Tennis Players." *Journal of Motor Behavior* 37 (2): 164–75.
- Shim, Jaeho, Les G. Carlton, and Young-Hoo Kwon. 2006. "Perception of Kinematic Characteristics of Tennis Strokes for Anticipating Stroke Type and Direction." *Research Quarterly for Exercise and Sport* 77 (3): 326–39.
- Siegel, Susanna. 2006. "Subject and Object in the Contents of Visual Experience." *The Philosophical Review* 115 (3): 355–88.
- . 2010. *The Contents of Visual Experience*. Oxford: Oxford University Press.
- Siewert, Charles. 1998. *The Significance of Consciousness*. Princeton: Princeton University Press.
- Simons, Daniel J., and Christopher F. Chabris. 1999. "Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events." *Perception* 28 (9): 1059–74.

- Simons, Peter. 1982. "The Formalisation of Husserl's Theory of Wholes and Parts." In *Parts and Moments: Studies in Logic and Formal Ontology*, edited by Barry Smith, 111–59. München, Germany: Philosophia Verlag.
- . 1994. "Particulars in Particular Clothing: Three Trope Theories of Substance." *Philosophy and Phenomenological Research* 54 (3): 553–75.
- . 2000a. "Continuants and Occurrents, I." *Aristotelian Society Supplementary Volume* 74 (1): 59–75.
- . 2000b. "Identity through Time and Trope Bundles." *Topoi* 19 (2): 147–55.
- . 2005. "Events." In *The Oxford Handbook of Metaphysics*, edited by Michael J. Loux and Dean W. Zimmerman, 357–85. Oxford: Oxford University Press.
- Smith, A. D. 2001. "Perception and Belief." *Philosophy and Phenomenological Research* 62 (2): 283–309.
- . 2003. *Routledge Philosophy GuideBook to Husserl and the Cartesian Meditations*. New York: Routledge.
- . 2008. "Husserl and Externalism." *Synthese* 160 (3): 313–33.
- Smith, Barry. 1989. "Logic and the Sachverhalt." *The Monist* 72 (1): 52–69.
- Smith, David Woodruff, and Ronald McIntyre. 1982. *Husserl and Intentionality: A Study of Mind, Meaning, and Language*. Boston: D. Reidel.
- Smith, Joel. 2016. *Experiencing Phenomenology: An Introduction*. London: Routledge.
- Soteriou, Matthew. 2010. "Perceiving Events." *Philosophical Explorations* 13 (3): 223–41.
- Spelke, Elizabeth S. 1990. "Principles of Object Perception." *Cognitive Science* 14 (1): 29–56.
- Strawson, P. F. 1959. *Individuals: An Essay in Descriptive Metaphysics*. New York: Routledge.
- Stroud, Barry. 2015. "Perceptual Knowledge and the Primacy of Judgment." *Journal of the American Philosophical Association* 1 (3): 385–95.
- Taieb, Hamid. 2018. *Relational Intentionality: Brentano and the Aristotelian Tradition*. New York: Springer.
- Textor, Mark. 2016. "States of Affairs." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2016. Stanford, CA: Metaphysics Research Lab, Stanford University. Retrieved from <<https://plato.stanford.edu/archives/win2016/entries/states-of-affairs/>>.



- Thompson, Brad J. 2009. "Senses for Senses." *Australasian Journal of Philosophy* 87 (1): 99 – 117.
- Thomson, Judith Jarvis. 1983. "Parthood and Identity Across Time." *The Journal of Philosophy* 80 (4): 201–20.
- . 1998. "The Statue and the Clay." *Noûs* 32 (2): 149–73.
- Tsai, Kellee S. 2007. *Capitalism without Democracy: The Private Sector in Contemporary China*. Ithaca, NY: Cornell University Press.
- Tye, Michael. 1991. *The Imagery Debate*. Cambridge, MA: The MIT Press.
- . 2014. "What Is the Content of a Hallucinatory Experience?" In *Does Perception Have Content?*, edited by Berit Brogaard, 291–308. Oxford: Oxford University Press.
- Van Brakel, J. 1996. "Interdiscourse or Supervenience Relations: The Primacy of the Manifest Image." *Synthese* 106 (2): 253–97.
- Vendler, Zeno. 1957. "Verbs and Times." *The Philosophical Review* 66 (2): 143–60.
- Ward, Paul, A. Mark Williams, and Simon J. Bennett. 2002. "Visual Search and Biological Motion Perception in Tennis." *Research Quarterly for Exercise and Sport* 73 (1): 107–12.
- Westerhoff, Jan. 2005. "Logical Relations between Pictures." *Journal of Philosophy* 102 (12): 603–23.
- Wolfe, Jeremy M. 2014. "Approaches to Visual Search: Feature Integration Theory and Guided Search." In *The Oxford Handbook of Attention*, edited by Kia Nobre and Sabine Kastner, 11–55. Oxford: Oxford University Press.
- Yablo, Stephen. 1992. "Mental Causation." *The Philosophical Review* 101 (2): 245–80.
- Yoshimi, Jeffrey. 2009. "Husserl's Theory of Belief and the Heideggerian Critique." *Husserl Studies* 25 (2): 121–40.
- . 2016. *Husserlian Phenomenology: A Unifying Interpretation*. New York: Springer.
- Yoshimi, Jeffrey, and David W. Vinson. 2015. "Extending Gurwitsch's Field Theory of Consciousness." *Consciousness and Cognition* 34: 104–23.
- Zacks, Jeffrey M., Nicole K. Speer, Khena M. Swallow, Todd S. Braver, and Jeremy R. Reynolds. 2007. "Event Perception: A Mind-Brain Perspective." *Psychological Bulletin* 133 (2): 273–93.
- Zacks, Jeffrey M., and Barbara Tversky. 2001. "Event Structure in Perception and Conception." *Psychological Bulletin* 127 (1): 3–21.

VITA

