

## University of Groningen

### Understanding for a purpose

Sleeuw, Stefan

DOI:  
[10.33612/diss.568507019](https://doi.org/10.33612/diss.568507019)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2023

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*  
Sleeuw, S. (2023). *Understanding for a purpose*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. <https://doi.org/10.33612/diss.568507019>

#### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

#### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



university of  
 groningen

# Understanding for a purpose

## PhD thesis

to obtain the degree of PhD at the  
University of Groningen  
on the authority of the  
Rector Magnificus Prof. C. Wijmenga  
and in accordance with  
the decision by the College of Deans.

This thesis will be defended in public on

Thursday 2 March 2023 at 11.00 hours

by

**Stefan Laurens Sleeuw**

born on 8 November 1990  
in Sneek

**Supervisors**

Prof. A.J.M. Peijnenburg

Prof. J.W. Romeijn

**Assessment Committee**

Prof. C.Z. Elgin

Prof. H.W. de Regt

Prof. M. Lenz

# Contents

Acknowledgements	1
Introduction	3
<b>1 The study of understanding</b>	<b>7</b>
1.1 Varieties of understanding	7
1.2 Historical background: epistemology	15
1.3 Historical background: philosophy of science	20
1.4 Introducing the aims of this project	26
1.5 Specifying the aims of this project	32
1.6 Conclusion	43
<b>2 Agents, signs, and symbols</b>	<b>45</b>
2.1 Agency	46
2.2 Peirce's concept of the sign	52
2.3 Subjective and objective signification	58
2.4 Type- and token-signs	63
2.5 Symbols and replicative signs	69
2.6 Conclusion	75
<b>3 Symbolic understanding</b>	<b>77</b>
3.1 How to use symbols	78
3.2 How to use symbols properly	86
3.3 Between propriety and competence	96
3.4 How to use symbols competently	101
3.5 Interlude: the nature of selective attention	109
3.6 Degrees of symbolic understanding: breadth	116
3.7 Degrees of symbolic understanding: depth	128
3.8 Conclusion	135
<b>4 Factual understanding</b>	<b>137</b>
4.1. Towards an explication of factual understanding	137
4.2. Explicating factual understanding	147
4.3. Clarifications	157
4.4. Replies to objections	162
4.5. Benefits of the explicatum	173
4.6 Conclusion	189
Concluding remarks	191
Nederlandse samenvatting	195
References	199



## Acknowledgements

This thesis has been written in the final 2 ½ years of my PhD candidacy. Almost all the text produced prior to that period has not made it into this book, and now sits somewhere, idly, in my file directory, perhaps to be unearthed in a distant future. The early documents bear remarkably little resemblance to anything in this dissertation. But at the same time, compiling those documents, and discarding them, has been absolutely vital to finding the research avenue that ultimately proved to be most fruitful. It is thanks to the formidable patience, trust, and enthusiasm of my supervisors, Jeanne and Jan-Willem, that I have been able to go through this somewhat roundabout but ultimately necessary selection process. Our many pleasant and lively conversations helped me greatly to stay motivated – periodically almost blissfully so – throughout the project. In addition to my supervisors, I would also like to thank Dr. Thomas Short, who devoted much of his scarce time to answering my annoying queries about his stunning book *Peirce's Theory of Signs* – the reading of which was instrumental to putting me on a new and more productive research path. I am similarly grateful to Prof. Catherine Elgin, who went through the trouble of reading, and commenting on, what was in hindsight a very peculiar section of my early work.

More than anything, the past four years have been filled with *thought*. Looking back, it seems as if almost everything in my life contributed in some way to creating an environment suitable for grim reasoning, utilitarian strategizing and unhinged speculation. The splendid isolation of a desk at the far end of a late-medieval attic; the serene regularity of a daily 15-minute bike commute; the steady (if occasionally also hectic) pace in my weekly round of package delivery through Rolde, Gasteren, Eext, Ekehaar, Schoonloo. All this tranquility would have quickly turned to chilliness, however, were it not for the warm blanket which the Groningen philosophy faculty is. During my time as a PhD candidate, I was fortunate to be able to participate in, and later preside over, the ever enjoyable monthly 'PCCP' meetings of the department of theoretical philosophy. A special thanks to the frequent contributors to the Epistemology reading group – most recently Job de Grefte, Chloé de Canson, Felix Friedrich – for the many interesting discussions we have had; discussions which often impacted on the course of my research. In addition, I wish to thank Janine Weeting for her invaluable support during times of indecision – the philosophy PhD community is very lucky to have her as an advisor.

My set of close (former) colleagues overlaps with my set of close friends to the point of equivalence. Regardless of set membership, therefore, I want to thank Merel, Hylke, Kritika, Jesse, Harm, Jochem, Justin, Reina, César and Caroline, for everything ranging from copious dinners to late-night role playing games at a campsite in Drenthe; from evening walks through Helpman to road trips in the beloved Škoda. Furthermore, I would like to highlight one element, not in the set of

## Acknowledgements

colleagues, which also happens to be the unique element in a very important singleton. Emma, thank you for being the person who (1) prevented us from entering Wok Bolsward, (2) co-conspired to acquire a marvelous piano and (3) is the best. Ten slotte, Johan, mocht je dit lezen: bedankt dat je er door dik en dun voor mijn moeder was, én dat je deur nog altijd openstaat.

Stefan Sleeuw, August 2022

## Introduction

During my time as a PhD student, I have enjoyed the privilege of being able to describe my research very succinctly, without *really* stretching the truth all that far. To a lay-person's inquiries about my work, I could often be heard answering along the following lines: "Many philosophers want to define what it means to know. I, on the other hand, am trying to define what it means to *understand*." And such a statement would be quite on point. I was, after all, trying to define what it means to understand. To my further advantage, I could even offer a fashionably provocative reply to the (relatively rare) follow-up question "well, what do you think it means to understand?" My go-to answer would be something like: "Many people are inclined to think that understanding happens in our brains or minds. I, however, believe that understanding consists in things we *do*." Unfortunately, I have had a much harder time justifying this belief in equally concise terms. The roguish "are you prepared to read 200 pages of dreary analytic philosophy?" may have saved me from alienating my conversation partners a couple of times, but such a cop-out move is of course no way to introduce those actual 200 pages to the reader.

So here I go. I believe understanding consists in behavior because we ascribe understanding to one another in order to regulate each other's doings. To ascribe understanding to someone is to encourage that person to continue manifesting certain behavior. Or, more accurately, it is to motivate that person to sustain a particular behavioral *disposition*. In this thesis, I flesh out this claim with respect to two subtypes of understanding, which I call *symbolic* and *factual understanding*. The former type pertains to our understanding of words, gestures, sentences, diagrams – or, in short, anything which has a representational function. The latter type concerns the understanding we may have of all kinds of non-representational entities: natural phenomena, historical events, persons, subject matters, etc. The aim of this dissertation is to explicate both these concepts in accordance with the general function I take understanding-ascriptions to have: the promotion of behavioral dispositions.

Why go through all this trouble? The immediate motive for this undertaking is dissatisfaction. Dissatisfaction, that is, with a debate about the nature of understanding that has been going on within epistemology for roughly two decades. My qualms revolve around the fact that the current discussion does not seem to have yielded results which properly answer to the concerns that first sparked the debate. Those concerns were quite far-reaching. According to some of the pioneers of 'the turn to understanding', the epistemological tradition had for too long placed undue weight on resolving largely technical problems within the analysis of knowledge, and had thereby lost sight of broader questions surrounding the importance of truth, the nature of epistemic value and the relevance of other cognitive (and behavioral) states besides belief. Their suggestion was that epistemologists would do well to



## Introduction

redirect some of their intellectual efforts to a different concept by means of which the broader questions could be approached in a fresh way: understanding.

This proposal struck a chord among a considerable number of epistemologists, but the revisionistic spirit that permeated some of the early proposals quickly watered down. Instead of taking the opportunity to reconsider fundamental (methodological) assumptions in targeting the novel (if also ancient) concept of understanding, many of the contributors to the understanding literature have stayed close to the knowledge-oriented epistemological tradition, to the point where part of the understanding debate has come to revolve around issues that were already central to the analysis of knowledge, most notably the problem of epistemic luck. To be sure, I am by no means claiming that such endeavors are altogether fruitless. I do, however, have a strong suspicion that the currently dominant approach is not likely to lead epistemology into the new territory that the early proponents of the turn to understanding had envisioned.

My own contribution to the understanding debate is meant to lie in how it repositions the study of understanding as a full-bodied alternative to the study of knowledge – an alternative which is equipped to deal with the broader questions just mentioned. This repositioning is effected by a change in methodology: it is epistemology's standard method of relying on intuitive judgments which I take issue with. Rather than taking intuitions as paramount, I prioritize accounting for the function that understanding-ascriptions have within a community. This function need not align with what we ordinarily believe understanding to consist in. In fact, as heralded by my fashionably provocative statement, I argue that the function-based approach prompts me to make some strongly counterintuitive claims. Those claims are not wholly without precedent, however. Although I am reluctant to label my proposal as 'Wittgensteinian' given the variety of possible interpretations of his musings about understanding in the *Philosophical Investigations*, one might discern a Wittgensteinian element in my persistent focus on the *practice* of ascribing understanding. A further, and much more salient influence is the philosophy of Charles Sanders Peirce. My account of the nature of symbols derives directly from Peirce's work, as does my externalist account of purposefulness, which underlies all occurrences of terms as 'function', 'goal', 'usage', and 'agency' throughout this dissertation. Although I again hesitate to call my proposal 'Peircean' for fear of misrepresentation, I would be happy to call it 'pragmatist'. Indeed, I would advertise it as such.

Let me finish this introduction by outlining the four chapters of which this dissertation consists. In **chapter 1**, I first disambiguate the term understanding, and delineate the two concepts that will take center stage in my inquiry: symbolic and factual understanding. Then, I survey the recent literature on understanding in epistemology and philosophy of science, and subsequently position my own views as regards the study of understanding in relation to this literature. In the final section, I offer a sneak peek into my accounts of symbolic and factual understanding, and

explain how I intend to defend these accounts. As I make clear, my general argumentative strategy relies on the format of *Carnapian explication*. This means that the fruitfulness of the proposed accounts of symbolic and factual understanding is meant to compensate for their intuitive deficits – intuitive deficits that result from the adoption of the function-based approach characterized previously.

**Chapter 2** lays important groundwork for the main investigation. For one thing, it provides the externalist account of the nature of purposefulness that undergirds every subsequent appeal to teleological notions in the dissertation. In addition, it demarcates the category of symbols within the larger category of *signs*, thereby setting the stage for the explication symbolic understanding. My proposal regarding the nature of symbols is thoroughly indebted to a recent reconstruction of Peirce's theory of signs due to Thomas Short. Complex and rich as Peirce's theory is, several sections are needed before I can formulate a definition of symbols. Given that the conceptual machinery introduced in those sections will be utilized frequently in chapter 3, however, their contents are also relevant to the dissertation as a whole.

Once the definition of symbols is in place, I turn to the task of explicating symbolic understanding. This, then, is what **chapter 3** is devoted to. Starting from the function-based approach, I argue over the course of four sections that ascriptions of symbolic understanding must pertain to the competent *usage* of symbols. In the final part of the chapter, I add to this proposal an account of how symbolic understanding can vary by degree along two axes: breadth and depth. Although the account of symbolic understanding can be considered an independent result of this inquiry, it ultimately performs an instrumental role within my dual explicatory project. That is, the outcome of the first explication (that of symbolic understanding) is meant to feature as an element within the outcome of the second: the explication of factual understanding.

It is in **chapter 4** that I develop and defend my account of factual understanding. Again, I argue that my preferred function-based approach occasions a behaviorist conception of this concept. More specifically, I submit that ascriptions of factual understanding must pertain to dispositions for transmitting useful information about the environment in which members of a community are situated. Having presented this view, I spend the second (and largest) part of the chapter resolving unclarity, countering potential objections, and, most importantly, highlighting benefits of my account. Among other things, those benefits concern the issues that initiated the turn to understanding in epistemology. I hope to show that my explication of factual understanding offers fruitful answers to those, and other, issues. Per Carnap's recipe, this serves to justify both my departure from our pretheoretic notion of factual understanding, as well as, indirectly, my departure from our pretheoretic notion of symbolic understanding.

## Introduction

### *Bibliographical note*

On a couple of occasions, I quote from the writings of Peirce. In doing so, I follow the established convention of using the format (*CP* x.y) to refer to a particular paragraph (y) within a volume (x) of the *Collected Papers*. Similarly, I use (*EP* x.y) to refer to a page (y) within a volume (x) of the *Essential Peirce*.

## Chapter 1: The study of understanding

Without understanding, life becomes very difficult. We are likely to be denied access to various primary necessities of life if we do not understand the non-verbal and verbal signals transmitted by the people in our vicinity. We must understand the workings of natural objects and artefacts in order to manipulate them in useful ways, or simply to avoid the potential dangers they pose. And we have to understand the lay-out of our environment so as to navigate through it successfully. Obviously, we can, and often do rely on others for social inclusion, protection, sustenance and movement. But in doing so, we are bound to rely heavily on *their* understanding of the world. Understanding being thus crucial to our survival and well-being, we might want to know – indeed, understand – what it amounts to. There are, however, some preliminary issues that have to be sorted out in preparation of taking on this rather grand question. First of all, understanding appears to be divisible into a number of subtypes. For instance, one may understand how to do something, why something happened, and what someone said. I therefore have to draw up a taxonomy of understanding, and subsequently specify which element in that taxonomy I shall focus on. Once this is done, I must make clear at what sort of result my investigation is aimed, and how that result is meant to be attained. This introductory chapter will be devoted mainly to addressing these two preliminaries: to charting the landscape of understanding and to laying out the methodology and goals of this project.

A brief outline of what follows is in order. In **section 1**, I take up the first task mentioned above. In doing so, I demarcate four main (partly overlapping) categories: propositional, explanatory, factual, and symbolic understanding. The latter two categories jointly constitute the main subject matter of this investigation. **Sections 2 and 3** subsequently discuss some of the main problems and ideas concerning understanding that have occupied epistemologists and philosophers of science. In **section 4**, I turn to the methodology and aims of this project. As I will argue, the method of conceptual analysis, of which many epistemologists continue to avail themselves, is less than optimal for studying understanding. Instead, I propose that we opt for Carnap's method of explication, and apply that method to the concepts of symbolic and factual understanding. **Section 5**, finally, will clarify what exactly it is that needs to be explicated and comment on how I will attempt to live up to the requirements associated with the method of Carnapian explication.

### 1.1 Varieties of understanding

The verb 'to understand' pervades the English language. So do its counterparts 'comprendre' and 'verstehen' in French and German, respectively. It is only since relatively recently, however, that epistemologists and philosophers of science have begun to seriously consider understanding as a concept worthy of philosophical study. In sections 2 and 3 of this chapter, I discuss in more detail some of the fruits

## 1 The study of understanding

of this burgeoning area of investigation. In the present section, I shall disambiguate the various ways in which we use the verb ‘to understand’ (and some of its cognates) in everyday speech, in an effort to acquire a more fine-grained picture of understanding qua philosophical concept. Thus, I start out by drawing some purely grammatical distinctions, and then move on to isolating a number of epistemological categories on the basis of these distinctions. Apart from facilitating the historical discussions in the ensuing sections, this preparatory exercise will allow me to specify the intended targets of this investigation: symbolic and factual understanding. As a disclaimer, I should note that there is an inevitable degree of arbitrariness with respect to the way in which the landscape of understanding is carved up in this section. Some might argue that my proposal misses out on genuine conceptual differences by relying on a merely superficial uniformity in common usage. Conversely, others may object that I all too readily translate differences identified at the level of ordinary language into full-blown conceptual distinctions. Nonetheless, it is evident that conceptual clarification has to start somewhere, and it seems to me that observations about the everyday use of a term provide an obvious basis for this, if by no means an infallible one.

The meaning of the word ‘understanding’ is blurry to say the least. Zagzebski (2009, p. 141-42), offers the following diagnosis of this semantic predicament:

Epistemology has been dominated by the values of certainty and understanding at different times in the history of philosophy, and the difference was reflected in the way knowledge was understood. As a rough generalization, the dominant value was certainty in eras marked by the fear of skepticism, and in those periods knowledge was closely associated with justification, since justification is what we want to defend our right to be sure. In contrast, understanding was the dominant value in those eras in which skepticism was not seen as threatening. Knowledge at those times was closely associated with explanation, since understanding is exhibited by giving an explanation. (...)

[S]kepticism has had an enormous impact on modern philosophy, so it is not surprising that understanding has received little attention. One of the sad consequences of neglect of a value is fragmentation of meaning. People can mean so many different things by the word ‘understanding’ that it is hard to identify the state that has been ignored. This can generate a vicious circle since neglect leads to fragmentation of meaning, which seems to justify further neglect and further fragmentation until eventually a concept can disappear entirely.

Regardless of whether this process of fragmentation will eventually be reversed through increased study, such study can only ever become fruitful after an analysis of the fragmentation itself. Let us therefore now try to reveal the forest behind the trees by means of some basic linguistic concepts. Despite rife debate over definitions and boundary cases, linguists typically distinguish between *monosemy*, *homonymy*, and *polysemy*. A term is said to be monosemous in case it has a single meaning,

homonymous when it has two or more unrelated meanings, and polysemous if it possesses two or more related meanings (Vicente & Falkum, 2017).<sup>1</sup> Technical and natural kind terms are usually monosemous: ‘positron’ and ‘sparrowhawk’ have just one meaning. An oft-used example of homonymy is the word ‘bank’, which can either stand for a river side or a financial institution. An instance of polysemy would be the word ‘dissertation’, which may be interpreted as a physical object, or as the totality of that object’s contents.<sup>2</sup> These two examples also convey in what sense homonymy differs from polysemy: the two senses of bank have little to do with each other, whereas the two senses of dissertation are clearly semantically connected.<sup>3</sup>

Natural language is rife with polysemy. The phenomenon is not limited to nouns and noun phrases; verbs can also have multiple, related readings. ‘To drink’, for instance, can refer to the act of drinking any substance, or to the consumption of alcohol. In this particular case, where one meaning is a specification of the other, the two senses are said to be *linearly* connected (Cruse, 2000, p. 110). In other cases, the relation between the several readings is not hierarchical in this way, and is thus characterized as *non-linear*. The verb ‘to get’, for example, has interpretations that bear only an abstract or metaphorical similarity to one another: one can get a prize, get pregnant, or get the punch line of a joke. The verb ‘to understand’ can also be taken to belong to the domain of such non-linear polysemous verbs: while not as far apart as the different senses of ‘to get’, ‘to understand’ has intuitively distinguishable, semantically non-hierarchically related senses.<sup>4</sup> To see this,

---

<sup>1</sup> I am here avoiding the perhaps more familiar notion of *ambiguity*, since this is in fact an umbrella term for various kinds of equivocation; including not only polysemy and homonymy, but also multi-interpretability at the syntactic level.

<sup>2</sup> That ‘dissertation’ is not monosemous is also demonstrated by the fact that a sentence such as ‘Judy’s dissertation is thought-provoking and yellowed with age’ sounds awkward (example taken from Norrick, 1981, p. 115).

<sup>3</sup> Although it has to be added that there is no clear-cut test for distinguishing polysemy from homonymy. One suggested criterion is etymological kinship: if the several meanings of a term have a common etymological origin, then that term is polysemous rather than homonymous. But this seems to relegate too many intuitive cases of homonymy to the domain of polysemy: even the two meanings of ‘bank’, for instance, have a common etymological origin in a distant linguistic past (see ‘bank’ in the *Online Etymology Dictionary*). Instead of searching for a more sophisticated test, we might as well conclude that strict monosemy and strict homonymy occupy opposing ends of a continuum, with polysemy covering all cases that fall somewhere in between these extremes.

<sup>4</sup> For the sake of completeness, I should note that I here only consider contexts in which ‘to understand’ is used as a transitive verb: as a verb that comes with a grammatical object. Sometimes, ‘to understand’ is used intransitively, such as in case of replies indicating comprehension or sympathy (e.g. ‘Of course I understand’). We may cautiously assume, however, that whatever analyses apply to the transitive variety also apply to the intransitive form. This is because the intransitive use of ‘to understand’ is a mere case of *object deletion*: the omission of a verb’s grammatical object for reasons of communicative expediency. Arguably, therefore, the distinction between transitive and intransitive uses of ‘to understand’ does not track any difference in meaning in the way that the distinction between the transitive

## 1 The study of understanding

consider the fact that someone may (fail to) understand Swahili, a close friend, the *Tao Te Ching*, quantum mechanics, how to cast a rod, why Pisa's tower is tilted, why his boss is angry with him, the meaning of life, and the meaning of this particular sentence. The senses of 'to understand' in these instances are closely related, yet it seems we cannot simply lump them together. In other words, not all these occurrences of 'to understand' are intuitively associated with the exact same kind of (mental) state or activity.

Some order can be imposed by categorizing different varieties of understanding in accordance with to the kind of item that 'to understand' can take as its grammatical object. First of all, the object can be a noun or noun phrase, such as 'a close friend' and 'the meaning of life'. But it can also be an interrogative clause that starts with an 'how', 'why', 'which', 'when' or 'what'. One may understand how the accident happened, why the relationship ended, which measures are appropriate, when to stop talking or what is expected during an emergency. Thirdly, the object of 'to understand' can be a declarative clause starting with 'that'. I may understand that the road is closed, or that my essay is long overdue. This use of 'to understand' is often thought to be synonymous with 'to know that' (Grimm, 2010, p. 85). Sometimes, however, understanding-that carries a subtle implication: if I utter a statement such as 'I understand that the chair of the committee will be arriving late today', this can likely be paraphrased along the lines of 'to my knowledge ...' or 'as far as I know ...'; indicating that I am hedging for reasons of doubt.<sup>5</sup> Finally, there is the possibility of 'to understand' being followed by 'how to' plus an infinitive. Thus one may understand how to pilot an aircraft, or how to make *beurre blanc*. Generally speaking, using this form amounts to ascribing an ability for successful performance to the subject: if I profess to understand how to fly an aircraft, you may reasonably expect me to be able to actually operate one.<sup>6</sup>

As per the disclaimer in the introduction to this section, we should be careful in translating these observations about common usage into philosophical categories straightaway. First of all, instances of distinct linguistic varieties may coincide on an epistemological level. For instance, what is expressed by means of an interrogative clause can usually also be expressed with a noun phrase: there is a clear semantic

---

and intransitive forms of, for instance, 'to see' does. See also Liu (2008) for an extensive discussion concerning the classification of intransitive verbs in English.

<sup>5</sup> See Gordon (2012) for a more detailed treatment of hedging uses of 'to understand that'.

<sup>6</sup> I should point out that most examples mentioned in this paragraph can also be expressed by means of the verbal noun 'understanding'. For example, one can have an understanding of how the accident happened, or possess an understanding of how to make *beurre blanc*. Admittedly, not all of the sentences so constructible sound equally natural. In addition, the noun form can also be used to indicate the existence of (implicit) agreements ('We quickly came to an understanding about the matter'), expectations ('According to my understanding, you would do the dishes') or mutual affection ('There is an intimate understanding between them'). To avoid complicating matters unnecessarily, however, I will confine my attention to the verbal form.

overlap (if perhaps not equivalence) between ‘Susan understands why Stalin rose to power’, and ‘Susan understands Stalin’s rise to power’. Either of these expressions can be used to refer to what is now commonly called *explanatory understanding*, which is the kind of understanding that is had by virtue of believing or being able to adduce sound explanations for events or phenomena. The epistemological category of explanatory understanding therefore does not coincide with any one grammatical category. A second ground for exercising caution in mapping the linguistic onto the philosophical, is that grammatically kindred expressions are often at odds with one another epistemically speaking. For example, instances of the form ‘to understand + interrogative clause’ need not always fall under the same epistemological heading. To understand *where* the nearest exit is, is a kind of epistemic state or achievement that is clearly different from the one involved in understanding *why* the nearest exit is located where it is.<sup>7</sup> What is more, there may even be epistemic discrepancies among expressions that are grammatically *identical*. As already mentioned, ‘to understand that’ is often interchangeable with ‘to know that’, but not in cases where ‘to understand that’ is used to convey doxastic reservations. Insofar as epistemologists recognize a category of *propositional understanding* (i.e. understanding that something is the case) this usually includes only the non-hedging variety.

Because of such mismatches, observations about ordinary language cannot be the sole determinant in carving up the landscape of understanding. Their role is rather heuristic: because we may expect that common usage reflects our philosophical predilections to at least some extent, identifying differences in common usage can serve as a starting point for further philosophical analysis. Here, I will take the liberty to single out the form ‘to understand + noun (phrase)’ as the focal point of attention, and examine it from an epistemological standpoint. For ease of reference, I will from now on refer to this category by the label of ‘objectual understanding’ – using quotation marks instead of italics in order to indicate that this is still only a grammatical category. The reason why I confine my attention to objectual understanding is that it is a very comprehensive category, but also a relatively understudied one. As Baumberger & Brun (2017, p. 165) explain:

Recent debates about the nature of understanding have mostly been focused on explanatory rather than objectual understanding. As a consequence, not only have the conditions for ascribing objectual understanding not been analyzed sufficiently, but the very notion of objectual understanding has remained somewhat unclear.

---

<sup>7</sup> This is not to deny that in many cases, different interrogative pronouns can be used to refer to the same epistemology category. For instance, understanding *why* the Baltic cruise ferry Estonia sank, *how* it sank, and *what* caused the ship to sink, all come down to the same thing (namely explanatory understanding).



## 1 The study of understanding

While some, such as Kvanvig (2003, p. 191) have opted for a broad, grammatically delineated notion of objectual understanding, others, such as Khalifa (2017, p. 80), have employed a more restrictive conception according to which objectual understanding pertains only to the understanding of subject matters, or bodies of information. I follow Kvanvig's line here, but will ultimately diverge from him in identifying within the grammatical category of objectual understanding two epistemological categories: symbolic and factual understanding. More on this distinction below.

To illustrate just how vast the scope of objectual understanding is, let me start off by listing a number of nouns and noun phrases that 'to understand' can take as its grammatical complement:

- Humans and other animals ('Oscar understands his best friend')
- Physical items ('Ariana understands the labyrinth')
- Mental items ('Frank understands his neighbor's desire for seclusion')
- Natural phenomena ('Mary understands ocean acidification')
- Historical events ('Arthur understands Napoleon's defeat at Waterloo')
- Subject matters ('Sylvia understands dendrochronology')
- Theories ('Caroline understands Darwin's theory of evolution')
- Representations ('John understands the Venn diagram')
- Languages ('Philip understands Tagalog')
- Meanings ('Samantha understands the Morse signal')
- Narratives ('Michael understands *The Neverending Story*')
- Concepts ('Jeanne understands *akrasia*')
- Genres ('Harry understands urban fantasy')<sup>8</sup>

Notwithstanding its grammatical uniformity, the category of objectual understanding harbors considerable epistemological diversity. Consider, first of all, the understanding of natural phenomena. To all appearances, this kind of understanding is closely wedded to explanation: it is difficult to see how we could understand solar eclipses, tidal flow or volcanic eruptions if we did not possess at least some account of why or how these phenomena occur. In fact, some authors have gone as far as claiming that understanding of this sort is reducible to explanatory understanding. According to Khalifa (2017), for instance, everything of epistemological relevance figuring in understanding a phenomenon can be cast either in terms of directly grasping explanations, or in terms of possessing information conducive to building explanations.

Consider now the understanding one may have of a language. If someone inquires whether you understand German, for instance, this is usually to be interpreted as a

---

<sup>8</sup> Although the contents differ, I should perhaps add that some of the inspiration for compiling this list was drawn from Wilkenfeld (2013, p. 998).

question about your versatility and facility in passive comprehension and active communication. Linguistic understanding, so conceived, does not bear an obvious relationship to explanation. Of course, understanding how modern-day German developed out of Old High German, why it is genetically closer to Dutch than to Danish, and why certain prepositions go with certain grammatical cases rather than others, are respectable epistemic achievements in their own right. But the believing or adducing of such explanations is neither a necessary, nor a sufficient condition for the kind of competence usually associated with linguistic understanding. This also applies at the level of understanding specific linguistic items. Understanding a sentence may be thought of as involving a cognitive component (e.g. correctly believing that a sentence means such-and-such) and behavioral one, such as the ability to use it in the appropriate circumstances, and for the right purposes. But being able to explain why the sentence possesses the meaning it has seems to be supererogatory to understanding the sentence. By extension, what goes for our understanding of natural languages and linguistic items, also goes for our understanding of other symbol systems and the representational devices of which those systems consist. Understanding Morse code, like understanding German, seems to center around an ability to engage with it passively and actively, rather than around explanation. And as argued for instance by De Regt (2009, 2017), a similar insight applies to the understanding of (scientific) theories, which are comprised of both linguistic (statements) and non-linguistic (formulas, diagrams) components. According to De Regt's influential account, understanding a scientific theory comes down to being able to use it through constructing adequate models.

Wrapping up the above: the usual sense of understanding (some item within) a language or other symbol system, while being superficially on a par with understanding a natural phenomenon in virtue of a shared grammatical form, seems to come apart from the latter in terms of the criteria that govern its attribution. There is thus ample reason to divide the epistemologically hybrid class of objectual understanding into more homogeneous subcategories. Following Baumberger, Beisbart & Brun (2017, p. 5), I propose to group uses of 'to understand' that pertain to linguistic and non-linguistic symbols and symbol systems under the heading of *symbolic understanding*; reserving the notion of *factual understanding* for "understanding that goes beyond representations and stretches out to facts, to phenomena of the world" (idem, p. 6).<sup>9</sup>

With the distinction between symbolic and factual understanding in hand, it appears we can neatly sort the various types of nouns and noun phrases listed previously. Understanding a language is obviously a matter of symbolic understanding, while understanding natural phenomena clearly belongs to the realm

---

<sup>9</sup> Strictly speaking, the term 'factual understanding' only appears in Baumberger (2014). Baumberger et al. (2017) do not employ the term 'factual understanding', but instead restrict their use of the term 'objectual understanding' to include only the factual variety.

## 1 The study of understanding

of factual understanding. As for some of the other items in the list, however, classification is less than straightforward. For instance, is understanding dendrochronology an instance of symbolic, or of factual understanding? On the one hand, dendrochronology consists in a method of dating tree rings that involves certain symbolic devices. Then again, it is also a branch of science that, through the application of its methods, aims to accurately determine the age of tree samples and to offer adequate explanations for certain patterns within those samples. In this sense, understanding dendrochronology entails the presence of an epistemic standing that “stretches out to phenomena of the world”. It seems that either interpretation can take precedence over the other, depending on contextual factors.

By way of illustrating this point further, consider also the different ways in which we can think of understanding narratives and concepts. At one level, understanding these entities involves the due recognition and application of the symbolic items that denote them, or of which they are composed. But at another level, narratives and concepts are more akin to natural phenomena, requiring sound explanation or, at least, reasonable argumentation in order to be understood. In section 4.1, I resolve these classificatory issues, by replacing the admittedly vague characterization of factual understanding provided above, with a ‘negative’ characterization, according to which factual understanding has as its objects all instances of objectual understanding (i.e. ‘to understand + a noun phrase) that fall outside the domain of symbolic understanding. At present, however, it suffices that we have a reasonably clear idea of how symbolic and factual understanding come apart, and how they fit into the overall landscape of understanding.

To close of this section, let me summarize its main findings and indicate how I will proceed. I have moved from a purely ordinary-language based inquiry into uses of the verb ‘to understand’, to an attempt at marking off epistemological varieties of understanding. I first touched upon explanatory understanding, and then briefly mentioned propositional understanding. Both of these epistemological categories were said to correspond imperfectly to distinctions drawn at the ordinary-language level. I then homed in on the sense of ‘to understand’ that takes a noun or noun phrase as its complement, and found that that grammatical subtype, referred to as objectual understanding, contains two epistemic types: symbolic and factual understanding. In closing, I have hinted at the fact that the symbolic and factual dimensions of some object of understanding are not always easily distinguishable. Further on in this chapter, I will explain how I intend to achieve the main purpose of this project, which is to explicate the concepts of symbolic and factual understanding. Before doing so, however, I will in the next two sections digress into some of the recent and not-so-recent precedents to the current investigation. This will allow me to properly situate my own intended contribution in section 1.4.

### 1.2 Historical background: epistemology

Within epistemology, a number of authors who had traditionally been concerned mainly with the study of knowledge, have in recent years repurposed their expertise to unearth the criteria governing the previously neglected concept of understanding. In this section, I survey some of the factors that contributed to the recent growth of interest in understanding in epistemology. In doing so, I shall for now sidestep a closely related development in philosophy of science, which is a topic that I will discuss in section 1.3. My main reason for doing so is the following. Despite the significant overlap between the two fields, epistemologists and philosophers of science often differ with regard to the nature of their interest in, and their approach to, the study of understanding. Epistemologists, for instance, have been primarily concerned with the relation between understanding and knowledge. Philosophers of science, on the other hand, have tended to focus almost exclusively upon the relation between understanding and explanation. While the division should not be overstated, keeping in mind these different starting points may serve to ward off potential confusion. In addition, by treating the two disciplines separately I seek to maintain clarity about the fact that my investigation is meant as a contribution, first and foremost, to epistemology.

Although epistemologists have only relatively recently begun to seriously investigate the concept of understanding, the notion itself is by no means a recent invention. As Greco (2014, p. 285) argues, for instance, the Aristotelian conception of *episteme* as involving ‘knowledge of causes’, ties in quite naturally with the modern-day use of the term understanding.<sup>10</sup> Furthermore, the British empiricists (e.g. Locke, Berkeley, Hume) made ample reference to understanding, albeit mainly in its nowadays uncommon sense as an overarching cognitive faculty or process, rather than as an epistemic state. In addition, understanding is arguably the most accurate translation of the German *Verstehen*, a notion that was introduced by the German historian Droysen and later used by Wilhelm Dilthey to argue for the methodological autonomy of the humanities. When it comes to placing the present work within a concrete epistemological debate, however, we can confine our attention to publications from 1990 and later. To my knowledge, it is only since then that epistemologists have begun to systematically theorize about understanding.

To see why epistemologists have started to turn their attention to the concept of understanding, we have to look at the intellectual history preceding that development. Throughout the second half of the twentieth century, many epistemologists have been engaged in developing an account of knowledge in terms

---

<sup>10</sup> In his introduction to Aristotle, Lear (1988, p. 6) defends a similar view regarding the proper translation of *episteme*: “Although ‘to know’ is an adequate translation of the Greek ‘*eidenai*’, Aristotle used this term generically to cover various species of knowing. One of the species is ‘*epistasthai*’ (literally, to be in a state of having *episteme*) which has often been translated as ‘to know’ or ‘to have scientific knowledge’, but which ought to be translated as ‘to understand.’ For Aristotle says that we have *episteme* of a thing when we know its cause.”

of necessary and sufficient criteria.<sup>11</sup> Almost exclusively, the target of analysis has been one specific kind of knowledge: the knowing of propositions by individual epistemic agents. More concretely, therefore, we can say that epistemologists working on this problem have tried to answer the question under what conditions some subject *S* can be said to know that *p*, where *p* is a proposition. According to a standard proposal that often figures as the starting point in discussions on this topic, propositional knowledge of the sort just described requires the *truth* of *p*, a *belief* that *p* on the part of *S*, and the existence of a suitable ground such that *S*'s belief that *p* is *justified*. Taken together, these three conditions are deemed to be sufficient for *S* to know that *p*. On this account, often referred to as the 'justified true belief (JTB) account' or 'tripartite analysis' of knowledge, mere true belief is insufficient for knowledge.<sup>12</sup> The insufficiency of true belief for knowledge is thought to derive from the incompatibility between knowledge and luck. As the argument goes, true belief can in principle be arrived at through lucky guessing, but such lucky guesses intuitively fall short of knowledge. Genuine knowledge must involve a true belief that is had for the right reasons, or acquired via the right causes, and the justification condition is meant to cover this requirement.

As is well-known to most readers familiar with the canon of twentieth-century analytic philosophy, Edmund Gettier's famously short paper 'Is Justified True Belief Knowledge?' (1963) challenged the idea that the justification condition suffices to exclude all cases of epistemic luck. By means of two scenarios, Gettier argued that an epistemic agent can possess a true belief that is irreproachably justified, but which nevertheless does not constitute knowledge due to the belief being only coincidentally true. Stated more generally, Gettier intended to show that the truth, belief and justification conditions may be individually necessary, but are not jointly sufficient for knowledge.<sup>13</sup> Gettier's argument quickly caught on, and caused "a

---

<sup>11</sup> Such purely conceptual work is sometimes categorized under the somewhat derogatory heading 'mainstream epistemology'. Mainstream epistemology is often contrasted with formal epistemology, which employs logical and mathematical tools for modelling and clarifying specific epistemic phenomena. See also Hendricks (2006).

<sup>12</sup> This idea dates back to Plato's *Theaetetus*, where Socrates argues against the proposal that knowledge is nothing more than true belief. According to Socrates, true beliefs can be acquired through mere persuasion by a clever advocate, whereas knowledge must be taught (see §§200-201c). Plato is often credited with having been the first to defend a JTB-like account of knowledge. This is seemingly evidenced by a passage from the *Theaetetus*, in which Theaetetus suggests that knowledge is "true belief with an account" (§201d). In what follows, however, Socrates rejects three ways of cashing out the idea that knowledge is true belief plus an 'account' or *logos*. The dialogue eventually ends in an *aporia* of sorts (§210b).

<sup>13</sup> Although Gettier was the one to put this issue on the philosophical agenda, he was by no means the first to come up with the kind of counterexample that is central to the argument. Perhaps the most well-known precedent is Bertrand Russell's clock example (Russell, 1948, p. 170-171), which runs as follows: "[Suppose] [t]here is [a] man who looks at a clock which is not going, though he thinks it is, and who happens to look at it at the moment when it is right; this man acquires a true belief as to the time of day, but cannot be said to have knowledge."

flurry of philosophical activity by epistemologists attempting to revise the JTB theory” (Ichikawa & Steup, 2018).<sup>14</sup> Some of these attempts involved the addition of further conditions that were designed to fend off Gettier-style epistemic luck scenarios.<sup>15</sup> Other authors took the alternative approach of strengthening, modifying or replacing the justification condition.

One influential proposal that falls into the latter camp was developed by Robert Nozick (1981). As he argued, Gettier-style counterexamples can be kept at bay if we replace the justification condition by a pair of modal criteria, so as to admit as instances of knowledge only those beliefs that ‘track’ the truth across relevant counterfactual circumstances. Although Nozick’s analysis thereby provides a solution to the original Gettier problem, there are amended Gettier-style counterexamples that do seem to undermine Nozick’s account after all.<sup>16</sup> A more recent proposal that does not fit neatly into either of the two categories just mentioned, is due to Ernest Sosa (2007). Sosa’s account averts Gettier cases not by adding independent criteria or by tampering with the justification condition, but by building in a connection between a belief’s truth and the way in which it is justified. According to Sosa, knowledge is constituted by beliefs that are true *because* they were formed in an epistemically appropriate manner. Again, this proposal arguably deals effectively with Gettier’s original counterexamples, but according to some ultimately fails to provide a solution to more sophisticated versions.<sup>17</sup>

It is important to note that proposals such as Nozick’s and Sosa’s, while to some extent departing from the JTB account, nonetheless remain faithful to certain assumptions implicit in that view. One of these assumptions is that knowledge is susceptible to conceptual analysis in the first place. This means that knowledge can in principle be spelled out in terms of simpler and more fundamental concepts. In his seminal work *Knowledge and its Limits* (2000), Timothy Williamson took issue with this idea, arguing that there are significant theoretical benefits to taking ‘knowledge first’, that is, to regarding knowledge as a fundamental, unanalyzable epistemic state. A second presupposition contained in traditional approaches to analyzing knowledge concerns a certain kind of optimism. According to this optimism, the problem posed by Gettier-style counterexamples can eventually be overcome through developing

---

<sup>14</sup> It is doubtful, however, whether Gettier’s article should be regarded as a turning point in the sense of having overthrown a long-established and widely accepted view. As Plantinga (1993, p. 6) remarks: “According to the inherited lore of the epistemological tribe, the JTB account enjoyed the status of epistemological orthodoxy until 1963, when it was shattered by Edmund Gettier with his three-page paper ‘Is Justified True Belief Knowledge?’. (...) Of course, there is an interesting historical irony here: it isn’t easy to find many really explicit statements of a JTB analysis of knowledge prior to Gettier. It is almost as if a distinguished critic created a tradition in the very act of destroying it.”

<sup>15</sup> See, for instance, Goldman (1967) and Lehrer & Paxson (1969).

<sup>16</sup> I am thinking here of Kripke’s ‘fake barn country’ example, which was, according to Adams & Clarke (2005, p. 214n11), first presented in an unpublished lecture in the 1980s.

<sup>17</sup> See, for example, Kvanvig (2010).

more sophisticated criteria. Linda Zagzebski cast doubt on the tenability of this attitude in her article ‘The Inescapability of Gettier Problems’ (1994). As she observed, Gettier-cases can be constructed for any account of knowledge that takes “the relation between justification and truth [to be] close but not inviolable” (Zagzebski, 1994, p. 65). Lest we become trapped in an endless dialectic, we are forced to either tighten the connection between truth and justification, or to admit Gettier cases as genuine instances of knowledge after all. This raises a dilemma: while the latter option is clearly at odds with our intuitions, the former seems to rule out the realistic possibility that we can sometimes justifiably believe what is false.

In addition to rejecting the analyzability of knowledge or denying that the Gettier problem has a definitive solution, there is another avenue for departing from the starting points of the traditional project.<sup>18</sup> This avenue leads us to the main topic of this investigation: understanding. It concerns a revisionary approach to epistemology pioneered by three philosophers: Catherine Elgin (1996, 2017), Linda Zagzebski (1996, 2001), and Jonathan Kvanvig (2003). Each of these authors has offered different arguments for the general idea that the focal point of epistemological inquiry – or at least one of its central subject matters – should be understanding rather than (just) knowledge. I will here briefly sketch their arguments so as to bring out how Elgin, Zagzebski and Kvanvig respond to, and diverge from, the broadly Gettier-oriented tradition that I discussed before.

Elgin, first of all, has long defended a view according to which epistemology should be more accommodating to those cognitive achievements whose epistemic worth seems self-evident, such as our most well-confirmed and fruitful scientific theories. In addition, Elgin thinks epistemology must in principle be able to account for the ways in which metaphors, fictions and thought experiments can function epistemically, in the sciences as well as in the arts. As she argues, the generally held assumption that knowledge requires truth is difficult to reconcile with the fact that our most esteemed scientific theories rely heavily on models and idealizations that depart from the truth in some way or another. Relatedly, the truth requirement also renders knowledge ill-suited for making sense of the epistemic value of metaphors, fictions and thought experiments. Primarily for these reasons, Elgin maintains that we had best reorient epistemological inquiry towards a concept that hinges on a more lenient conception of epistemic ‘rightness’. This concept, unsurprisingly, is understanding.<sup>19</sup> In her recent book *True Enough* (2017), Elgin develops a broadly Kantian framework for explicating the kind of rightness required for understanding.

---

<sup>18</sup> I am in fact skipping over some further options here. One prominent alternative is to accept that knowledge is analyzable, but to reject the existence of a unique standard of epistemic evaluation. This is roughly the stance taken by epistemic contextualists. Some important contributions within this school include Lewis (1996) and DeRose (1995).

<sup>19</sup> A precedent to this view can be found in the works of Nelson Goodman (more specifically his *Languages of Art* (1968) and *Ways of Worldmaking* (1978)), with whom Elgin has collaborated intensively.

## 1.2 Historical background: epistemology

As said, Zagzebski (2001) has likewise promoted a reorientation of epistemological inquiry towards understanding. One of her arguments in favor of such a reorientation is based on historical considerations. According to Zagzebski, we should be aware that epistemology's preoccupation with knowledge is by no means a self-standing and unalterable given. As she points out,

[T]he questions that we consider central to epistemology change over time and are not presented for our reflection singly, but in clusters. An important way in which questions bunch together is around the issue of skepticism. There have been significant periods of philosophical history in which skepticism was thought to be a serious threat, and other periods in which it was not. In those eras in which it *was*, philosophers gave preeminence to the epistemic value of certainty and focused on the nature of justified belief. (Zagzebski, 2001, p. 236)

As Zagzebski goes on to argue, the contemporary focus on skeptical worries has for some time exerted a crippling influence on epistemology. In order to open up new perspectives, she suggests that we take our cue from those periods in which skepticism was not at the forefront of epistemologists' concerns. Zagzebski believes that Plato provides one such entry point. She relies on an interpretation due to Moravcsik (1979), according to which *epistêmê* does not pertain to propositions but rather to entire fields, and involves, or is intimately tied to, the notion of *technê*: a practical art or skill. Zagzebski (2019) has developed this idea into a theory of understanding that takes 'grasping structures of reality' to be constitutive of understanding.

The third and final author, Kvanvig, has approached the topic from yet a different vantage point. According to Kvanvig, what is perhaps most objectionable about 20<sup>th</sup>-century epistemology is the fact that it seems unable to deal with the problem of epistemic value. Kvanvig (2017, p. 181-182) identifies two main questions surrounding epistemic value. First, there is the general question why the presence of an epistemic state (e.g. knowledge) is usually more valuable than its absence (e.g. ignorance). Secondly, we may want to explain why some epistemic states are more valuable than others. One such comparative question regarding epistemic value is why the value of knowledge exceeds that of any of its parts. A precedent to this question can be found in Plato's *Meno*, where Socrates endeavors to explain why knowledge is more valuable than mere true belief. Kvanvig (2003) argues that the contemporary epistemologist should likewise explain why knowledge is more valuable than justified, but not Gettier-proof, true belief. This, however, he thinks is a question that traditional approaches cannot answer:

[A]s we have learned from Gettier (1963), knowledge is more than normatively appropriate true belief, and so to show that knowledge is more valuable than any combination of proper subparts, we would have to be able to show that unGettiered normatively appropriate true belief is preferable to Gettiered normatively appropriate



## 1 The study of understanding

true belief. Kvanvig [(2003)] examines the variety of approaches to the Gettier problem, and argues that in each case, the better they get at avoiding counterexamples, the worse they are in terms of resources for solving this special value problem. (Kvanvig, 2017, p. 183)

The central worry behind Kvanvig's argument is that an anti-Gettier condition, however construed, will likely be too gerrymandered to plausibly count as the bearer of knowledge's distinctive epistemic value. To avoid having to conclude that our primary epistemic good does not carry a distinctive epistemic value, we should instead place our bets on understanding, which, according to Kvanvig, has a value that clearly exceeds the value of knowledge.

The contributions of the three authors just mentioned are widely recognized as having been instrumental in reshaping epistemology's agenda. Their rekindling of the concept of understanding has resulted in a proliferation of research on that topic. To a significant extent, that research parallels the earlier work done on knowledge. For instance, a considerable number of authors think that understanding should come with criteria that are similar to the belief, truth and justification conditions on knowledge. In addition, there is a growing literature on the compatibility between understanding and various forms of epistemic luck.<sup>20</sup> In section 1.4, I will argue that the study of understanding is unlikely to benefit from this close mirroring of the traditional study of knowledge. If our aim is to avoid replicating the problems hinted at by Elgin, Zagzebski and Kvanvig, we must be prepared to reconsider our main methodological assumptions and to redefine or replace some basic concepts. The alternative I shall suggest in section 1.4, and which I will work out in this dissertation, is to *explicate* (in Carnap's sense), rather than *analyze*, the concept of understanding. More specifically, the targets of explication will be the concepts of symbolic and factual understanding. The idea is that by subjecting these concepts to the method of explication, we are in a better position to pursue the revisionary ideals cherished by the pioneers of understanding.

### 1.3 Historical background: philosophy of science

As indicated before, the revival of interest in understanding occurred in two philosophical disciplines: epistemology and philosophy of science. Although the debates on understanding within these two areas have been sufficiently self-contained to warrant separate treatment in this introductory chapter, it would be mistaken to say that there is no overlap or interaction at all. Epistemologists and philosophers of science alike aim to identify the criteria for understanding, and especially in recent years there has been more and more cross-fertilization regarding this issue.<sup>21</sup> What is distinctive about the study of understanding within philosophy of science, however, is that it is tightly linked to the study of scientific explanation.

---

<sup>20</sup> See, for instance, Pritchard (2014) and Carter, Pritchard & Shepherd (2019).

<sup>21</sup> See Grimm, Baumberger & Ammon (2017).

This section highlights some of the views concerning understanding and explanation that have been defended by philosophers of science from the second half of the 20<sup>th</sup> century onwards.<sup>22</sup> Such a survey is called for because, although the present project is meant first and foremost as a contribution to epistemology, one of my target concepts, factual understanding, is intuitively closely linked to explanatory understanding (see, for instance, my remarks on understanding natural phenomena in section 1.1). It is thus apt to acquaint ourselves with what philosophers of science have had to say about understanding and explanation, even if this dissertation does not aim to directly contribute to the literature on that subject.

A natural starting point for a discussion of understanding in philosophy of science is without question Carl Hempel's work on scientific explanation. In his article 'Studies in the Logic of Explanation' (1948), co-authored with Paul Oppenheim, Hempel put forward a formal characterization of explanation, which is variously known as the deductive-nomological (DN), or the covering law model of explanation.<sup>23</sup> According to Hempel, scientific explanation involves the subsumption of natural phenomena under natural laws according to the following argument scheme (Fetzer, 2017):

Premises:	$L_1, L_2, \dots L_k$
	$C_1, C_2, \dots C_r$
	-----
Conclusion:	$E$

In this scheme, a description of the phenomenon to be explained (called *explanandum*) takes the place of the argument's conclusion ( $E$ ). That which does the explaining (called *explanans*) is constituted by the set of premises, which consists of sentences expressing natural laws ( $L_1, L_2, \dots L_k$ ) and statements that describe initial conditions pertaining to the phenomenon ( $C_1, C_2, \dots C_r$ ). For an explanation of this sort to be adequate, the explanandum must be a logical consequence of the

---

<sup>22</sup> One of the merits of Henk de Regt's work *Understanding Scientific Understanding* (2017) is that it provides the reader with a comprehensive summary of the most important work on understanding in philosophy of science. The overview I present here (especially its chronology) relies heavily on De Regt's, to the extent that merely providing explicit citations would not suffice as an exhaustive representation of actual indebtedness. Hence, let me emphasize here that the structure of my treatment largely follows that of the work already done by De Regt.

<sup>23</sup> Hempel and Oppenheim's article starts out with the ostensibly commonplace assertion that "[t]o explain the phenomena in the world of our experience, to answer the question 'why?' rather than only the question 'what?' is one of the foremost objectives of all rational inquiry". (Hempel & Oppenheim, 1948, p. 135) As Dewulf (2018) has recently shown, however, the idea of explanation being an important aim of scientific inquiry was not yet widely accepted in Hempel's time. Sustained interest in explanation among philosophers of science seems to have emerged only after, and perhaps largely as a result of, Hempel's publication.

## 1 The study of understanding

explanans, the explanans must be empirically testable, and the sentences that make up the explanans must be true (Hempel, 1965, p. 247-248).

On Hempel's analysis, explanation is taken to consist in a logical relation between certain categories of statements. In this sense, it is objective: the adequacy of explanations does in no way depend on their being intelligible to individual subjects. Hempel acknowledges that explanation admits of a 'pragmatic interpretation', which "concerns a three-term relation between explanation, phenomenon, and the person who uses the explanation" (De Regt, 2017, p. 19). However, he argues that this pragmatic sense is not relevant for philosophy of science:

For scientific research seeks to account for empirical phenomena by means of laws and theories which are objective in the sense that their empirical implications and their evidential support are independent of what particular individuals happen to test or to apply them; and the explanations (...) based upon such laws and theories are meant to be objective in an analogous sense. (Hempel, 1965, p. 426, as cited in De Regt, 2017, p. 18)

Relatedly, insofar as Hempel makes mention of understanding, he only employs a conception closely connected to, if not synonymous with, his objective interpretation of explanation. On this reading, understanding is taken to be an "objective kind of insight that is achieved by a systematic unification, by exhibiting the phenomena as manifestations of common underlying structures and processes that conform to specific, testable, basic principles" (Hempel, 1966, p. 83, as cited in De Regt, 2017, p. 50). Despite the use of the term 'insight', Hempel's conception of understanding is far removed from the more subjective, psychological sense associated with ordinary-language use of the word understanding.<sup>24</sup>

With regard to the study of understanding in philosophy of science, Hempel's work has been singularly influential. This is evidenced, first of all, by the fact that philosophers of science almost unanimously think of understanding as something that is constituted by, or even reducible to, correct explanations. A further mark of Hempel's lasting influence is the fact that philosophers of science, unlike epistemologists, have mostly concentrated on an impersonal conception of

---

<sup>24</sup> While Hempel regards the study of pragmatic aspects of explanation and understanding as falling outside the purview of philosophy of science proper, he devotes some attention to discussing a view that does take such aspects into account. According to that view, "an explanation must somehow reduce or link the puzzling phenomenon to something with which the questioner is already familiar, and which he accepts as unproblematic" (Hempel, 1965, p. 430). Hempel dismisses the idea that explanation amounts to such a 'reduction to the familiar', by noting that scientists routinely explain seemingly familiar events and phenomena in less familiar terms: "instead of reducing the unfamiliar to the familiar, a scientific explanation will often do the opposite: it will explain familiar phenomena with the help of theoretical conceptions which may seem unfamiliar and even counter-intuitive, but which account for a wide variety of facts and are well supported by the results of scientific tests" (Hempel, 1965, p. 431).

understanding. That is, philosophers of science tend to conceive of understanding as an epistemic product of science as a whole, rather than as an epistemic state entertained by individual agents. A third Hempelian inheritance is arguably the relative popularity of a theory of explanation-based understanding known as *unificationism*. According to the view developed by Michael Friedman (1974, p. 15), for instance, “science increases our understanding of the world by reducing the total number of independent phenomena that we have to accept as ultimate or given”.<sup>25</sup> A proposal in the same spirit as Friedman’s is due to Philip Kitcher. As Kitcher (1989, p. 432) claims, “science advances our understanding of nature by showing us how to derive descriptions of many phenomena, using the same patterns of derivation again and again, and, in demonstrating this, it teaches us how to reduce the number of types of facts we have to accept as ultimate (or brute)”. While both Friedman and Kitcher take issue with certain aspects of Hempel’s view, they retain the idea that explanation essentially consists in the subsumption of descriptions of phenomena under laws.

This fundamental idea has also met with some notable objections, however. An important line of criticism has it that mere subsumption is far too permissive a criterion: many deductive arguments that satisfy the deductive-nomological model intuitively fail to be explanatory. A perhaps familiar example can serve to illustrate this. On Hempel’s account, the length of the shadow cast by a flagpole can be explained in terms of a natural law – the fact that light travels in straight lines – conjoined with particular facts – the flagpole’s height and the sun’s elevation above the horizon.<sup>26</sup> By the same token, however, the DN model allows us to explain the height of the flagpole in terms of the rectilinear propagation of light, the position of the sun, and the length of the shadow. While the former seems appropriate as an explanation, the latter does not.<sup>27</sup> As Michael Strevens (2008, p. 24) puts it, “[t]here

---

<sup>25</sup> As observed by De Regt (2017, p. 52), Friedman does not seem to be entirely aware of his Hempelian roots. On the one hand, he distances himself from that tradition by arguing that Hempel and others have been too quick in denying the philosophical relevance of a psychological notion of understanding. As Friedman (1974, p. 8) states: “[A]lthough the notion of understanding, like knowledge and belief but unlike truth, just is a psychological notion, I don’t see why it can’t be a perfectly objective one. I don’t see why there can’t be an objective or rational sense of ‘scientific understanding’, a sense on which what is scientifically comprehensible is constant for a relatively large class of people.” Then again, the unificationist account of understanding Friedman ends up defending in no way hinges upon psychological notions; it seems that Hempel could have readily subscribed to it.

<sup>26</sup> Provided, of course, that the explanandum is deductively entailed by the explanans.

<sup>27</sup> Hempel was in fact aware of this type of counterexample to the DN-model. He considers a scenario, similar to the flagpole case, in which a pendulum’s length is accounted for in terms of its period, and asks whether this amounts to a proper explanation of the pendulum’s length (Hempel, 1965, p. 353). Interestingly, Hempel did not think this example posed a significant problem for his theory:

The law for the simple pendulum makes it possible not only to infer the period of a pendulum from its length, but also conversely to infer its length from its period; in either case, the inference is of the form (D-N). Yet a sentence stating the length of a

is an asymmetry in explanatory power where there is no asymmetry in logical power”. According to some, this asymmetry suggests that explanation is inextricably connected to causation, rather than to unification. Causal relations, after all, are clearly asymmetrical: if *A* is a cause of *B*, then *B* is not a cause of *A*. Hence, to explain a phenomenon might amount to identifying what caused that phenomenon.

Wesley Salmon is among the most well-known proponents of such a causal account of explanation. For Salmon, explanations lay bare the causal history leading up to the occurrence of an event. Such causal histories consist of processes that are “capable of propagating marks or modifications imposed on them” (Salmon, 1998, p. 131).<sup>28</sup> Understanding, in turn, is achieved by “the development of a world-picture” that is couched in terms of such causal-process based explanations (idem, p. 90). Although it stands opposed to Hempel’s broadly unificationist framework, Salmon’s theory nonetheless remains loyal to two of the three Hempelian commitments mentioned before. First of all, it is apparent that for Salmon, understanding is constituted by the coalescence of well-confirmed explanations into a world-picture. Furthermore, he does not deal explicitly with understanding as an epistemic state of individuals, but relies on a notion of understanding that pertains to science, or the scientific community, as a whole.

Perhaps inspired by parallel developments in epistemology, more recent work on understanding within philosophy of science increasingly takes the subjective dimension into account, while staying close to an explanation-centered view of understanding. Strevens (2013, p. 510), for instance, asserts that “[a]n individual has

---

given pendulum, in conjunction with the law, will be much more readily regarded as explaining the pendulum’s period than a sentence stating the period, in conjunction with the law, would be considered as explaining the pendulum’s length. This distinction appears to reflect the idea that we might change the length of the pendulum at will and thus control its period as a “dependent variable,” whereas the reverse procedure does not seem possible. This conception is questionable, however; for we can also change the period of a given pendulum at will, namely, by changing its length. It cannot validly be argued that in the first case we have a change of length independently of a change of the period, for if the location of the pendulum remains fixed, then its length cannot be changed without also changing the period. In cases such as this, the common-sense conception of explanation appears to provide no clear grounds on which to decide whether a given argument that deductively subsumes an occurrence under laws is to qualify as an explanation.

<sup>28</sup> This criterion, which is based on the so-called *mark principle* developed in Hans Reichenbach’s *The Direction of Time* (1956), serves to differentiate causal processes from what Salmon calls ‘pseudo-processes’. As Salmon (1998, p. 131) illustrates, “[a]n automobile traveling along a road is an example of a causal process. If a fender is scraped as a result of a collision with a stone wall, the mark of that collision will be carried on by the car long after the interaction with the wall occurred. The shadow of a car moving along the shoulder is a pseudo-process. If it is deformed as it encounters a stone wall, it will immediately resume its former shape as soon as it passes by the wall. It will not transmit a mark or modification.”

### 1.3 Historical background: philosophy of science

scientific understanding of a phenomenon just in case they grasp a correct scientific explanation of that phenomenon.” As he goes on to explain,

The view does not reduce understanding to explanation – the psychology of grasping is important and far from trivial (...). But it does [imply] that explanation is essentially involved in scientific understanding, and that the norms of correct scientific explanation logically precede and participate in determining the nature and norms of understanding. (Ibidem)

Interestingly, Strevens characterizes grasping as being itself a form of understanding. As he is quick to point out, no circularity is present here, since the understanding required for grasping a true explanation is of a different variety than the understanding obtained by grasping a true explanation. The former is *understanding-that*, while the latter is *understanding-why*. As Strevens (idem, p. 511) argues, the propositional understanding involved in grasping, cannot be reduced to standard propositional knowledge: “Someone with relatively little understanding of chemistry can, I think, know that water is made up of H<sub>2</sub>O, or that mercury is a metal. But they do not thereby grasp that these states of affairs hold in the sense required for understanding the chemical properties of water or mercury.” Grasping, on Strevens’ account, “is the fundamental relation between mind and world, in virtue of which the mind has whatever familiarity it does with the way the world is” (ibidem).

Kareem Khalifa (2017) also attempts to reconcile the explanation-based conception of understanding prevalent in philosophy of science with an epistemological, subject-oriented approach. What is perhaps most distinctive about Khalifa’s account is that it takes a comparative conception of understanding to be the “fundamental unit of analysis” (Khalifa, 2017, p. 14). In other words, it offers a criterion for when a subject  $S_1$  has *better* understanding of the explanation of a phenomenon than a subject  $S_2$ , rather than a criterion that determines when a subject possesses understanding *simpliciter*. According to Khalifa, a subject’s explanatory understanding of a phenomenon is better than someone else’s, just in case her grasp of the ‘explanatory nexus’ of that phenomenon is stronger. The explanatory nexus consists in “the set of correct explanations of [the phenomenon] as well as the relations between those explanations” (Khalifa, 2017, p. 6). Grasp, as used here, is “nothing more than a cognitive state bearing some resemblance to scientific knowledge of some part of the explanatory nexus” (idem, p. 11). Scientific knowledge, in turn, is possessed just in case one’s belief concerning a certain explanation within the nexus was formed on the basis of an evidence-based, comparative evaluation of potential explanations. Given that understanding is here defined in terms of grasp, and grasp in terms of scientific knowledge, Khalifa’s

account effectively collapses into the view that explanatory understanding amounts to scientific knowledge of explanations.<sup>29</sup>

Henk de Regt's (2017) contextual theory of scientific understanding, finally, constitutes yet a different kind of departure from the Hempelian starting points. While continuing the traditional focus on explanatory understanding, De Regt distances himself from the idea that understanding, insofar as its philosophically relevant features are concerned, is more or less reducible to the concept of explanation. As he argues, explanation is only one part of the story. For in order to generate and evaluate explanations, scientists need to be able to construct models. And for this, in turn, scientists must possess understanding of the theories from which the models can be derived. This kind of understanding crucially requires cognitive skills which defy translation into explicit rules and which can only be acquired within, and assessed by, a particular scientific community (De Regt, 2017, p. 28-29). Scientific understanding, according to De Regt, is thus clearly irreducible to explanation because it has an ineliminable pragmatic dimension. Still, he does seem to adhere to the impersonal perspective on understanding that is implicit in most work on understanding within philosophy of science. In this sense, his approach is somewhat further removed from the epistemological tradition than that of Strevens and Khalifa.

Let me round off this section by explaining in a bit more detail how my own work connects to understanding-related research in philosophy of science. As mentioned earlier, this dissertation purports to contribute primarily to the epistemological debate on understanding. It does so by providing explications of the concepts of symbolic and factual understanding. As noted before, factual understanding appears to be intimately linked (if perhaps not reducible) to explanation: intuitively, for instance, natural phenomena and historical events must be explained in order to be understood. Whatever light can be shed on the nature of factual understanding might thus indirectly illuminate the relation between understanding and explanation. Furthermore, to the extent that symbolic understanding is necessary for, or instrumental to, factual understanding (a topic I discuss in chapter 4), my explication of the former concept may also bear on how we take explanation to relate to understanding. For these reasons, the contents of this dissertation are of potential interest not only to epistemologists, but also to philosophers of science engaged in the study of understanding and explanation.

#### **1.4 Introducing the aims of this project**

Within analytic philosophy, the dominant approach to answering philosophical questions has long consisted in identifying the necessary and sufficient conditions

---

<sup>29</sup> Khalifa sides with Hempel with regard to the philosophical relevance of understanding. As he puts it: “[G]ive me an epistemology of scientific explanation, and I will give you everything you could have wanted from a philosophy of understanding” (Khalifa, 2017, p. 20).

for a concept's applicability on the basis of reflection upon our intuitive judgments. Through using this method, epistemologists have tried to formulate definitions of knowledge and its concomitant concept of justification. Notwithstanding the recent 'turn to understanding', the dominant approach has remained largely intact. The current project, while latching on to recent developments in terms of the subject matter it addresses, takes issue with the methodological underpinnings of much contemporary work on understanding. The first part of this section presents two arguments for the claim that the standard method of conceptual analysis is not ideally suited for constructing an account of understanding. Drawing on Baumberger (2019), I will first of all argue that the scalar character of understanding-attributions is difficult to reconcile with the fact that necessary and sufficient conditions determine a concept's applicability in an all-or-nothing manner. Secondly, I shall explain why I think that the intrusion of knowledge-related concepts and problems into the study of understanding, which is what adherence to the standard approach is currently leading to, threatens to undermine the relevance of conducting that study. Having argued so, I will in the second part of the section explain what sort of alternative I intend to offer in this dissertation.

The recent epistemological literature on understanding mimics the post-Gettier literature on knowledge in several respects. For instance, we often find discussions of such as questions as: does understanding require belief? Is understanding factive? What sort of justification condition applies to understanding? And like the post-Gettier knowledge literature, such questions tend to be tackled by consulting our intuitions through examples and counterexamples. This traditional approach has a number of advantages. First of all, it provides the research programme with a firm footing: by remaining faithful to the time-honored approach, the understanding literature comes to occupy a recognizable position within the epistemological tradition. Relatedly, sticking with the standard *modus operandi* allows authors to forgo dreary preliminaries on terminological and methodological issues. Since philosophers trained in the analytic canon are accustomed to the use of notions such as belief, truth and justification, as well as to the case-driven style of argumentation, they can get straight to business. Thirdly, and perhaps most importantly, exercising methodological conservatism renders knowledge and understanding commensurable. That is, by applying roughly the same format to the study of knowledge and of understanding, the two concepts can be readily compared with one another.

Despite such benefits associated with the traditional approach, however, there are at least two downsides to it. One worry, voiced by Baumberger (2019, p. 369), is that "the debate [on understanding] is at risk of running idle", because of a neglected dissimilarity between knowledge and understanding. Following Gettier's impactful publication in 1963, a considerable portion of the epistemological literature has been dominated by the implicit belief that, in principle, the concept of knowledge admits of a definition: an analysis in terms of individually necessary and jointly sufficient



## 1 The study of understanding

criteria.<sup>30</sup> Many scholars have been guided by the assumption that such a definition can be extracted from careful reflection upon our intuitions. There is one respect in which the concept of knowledge lends itself well to this approach. This concerns the fact that we ordinarily think of propositional knowledge as an all-or-nothing affair: either one knows a proposition to be true, or one does not.<sup>31</sup> This binary aspect of our standard conception of knowledge sits well with the project of providing an analysis in terms of necessary and sufficient criteria. In case of understanding, however, the picture looks different. Unlike attributions of knowledge, ascriptions of understanding are often qualified in terms of relative strength. It is completely natural, for instance, to say that someone understands something thoroughly, reasonably well, or only superficially, or to say that someone's understanding is broader, deeper, or otherwise better than someone else's. If we take it that an epistemological account of understanding is to respect such features of ordinary discourse, trying to identify a set of individually necessary and jointly sufficient criteria for 'outright' understanding would seem an overly Procrustean enterprise. Baumberger (*ibidem*) expresses this concern as follows:

If we agree that a suggested condition admits of degrees and constitutes an evaluative criterion, disagreement about how well it must be met in a certain context for outright understanding is comparatively undramatic. Whether someone can in some context be credited with understanding a subject matter becomes less urgent if we agree that her understanding would at least be minimal, for instance because the inferences she can draw about the subject matter are very limited, or because her theory is inconsistent with her background beliefs.<sup>32</sup>

---

<sup>30</sup> As in section 1.2, I am employing the usual interpretation of knowledge as propositional knowledge.

<sup>31</sup> In *Good Knowledge, Bad Knowledge* (2001), Stephen Hetherington characterizes this 'epistemic absolutism' as a central dogma of contemporary epistemology. But see, for instance, Moss (2018) for an influential example of a view that deviates from this dogma.

<sup>32</sup> In spite of widespread recognition that understanding admits of degrees, many authors continue to take outright understanding as their focal point of analysis. In light of the aim of comparing knowledge and understanding, this makes some sense. Employing an outright conception of understanding makes it possible, for instance, to investigate whether knowledge and understanding differ with respect to their compatibility with epistemic luck. One could question, however, to what extent the results of such an investigation are reliable. For how can we be sure that the intuitions we take as input to our theorizing derive from a contextually robust binary conception of understanding, given that our pretheoretic judgments otherwise point to understanding being a matter of degree? In essence, the problem alluded to here is one of underdetermination: intuitions regarding the (in)compatibility of understanding with epistemic luck could in principle be due to the threshold level for outright understanding being (unwittingly) set at a particular level according to particular aspects of a sample case, rather than to a constitutive feature of understanding *per se*. More generally, this kind of underdetermination threatens to undermine any comparative, intuition-based investigation in which the goal is to assess whether a certain feature associated with knowledge can also be ascribed to understanding.

The moral to be drawn from these considerations is that it would be naïve to expect that a full-fledged conceptual analysis of outright understanding will simultaneously (1) answer to all our relevant intuitions and (2) provide universally binding criteria. Given the fact that understanding intuitively comes in degrees, any attempt at meeting (2) is likely to come at the expense of (1), and vice versa.

A second issue with perpetuating the traditional approach is that it is unlikely to yield insights that are interesting and worthwhile in light of what initially motivated the turn to understanding. As said, there has thus far been a tendency to look at the concept of understanding through knowledge-t(a)inted glasses, resulting, for instance, in a preoccupation with the question how the tripartite analysis (justified true belief) could be tweaked so as to procure a satisfactory definition of understanding. I already argued that the search for an all-or-nothing definition does not accord well with the graded character of understanding. But on top of that, there is the worry that the conservative project carries little promise beyond offering an afterthought, or footnote, to decades of theorizing about knowledge. Such a modest outcome would be at odds with the lofty ambitions that once inspired certain epistemologists to opt for a wholesale intellectual reorientation. Their ambitions sprang from dissatisfaction not just with a single concept, but rather with an entire research programme, or paradigm. Roughly put, that paradigm revolved around the quest for securing the link between our cognitive endeavors and the ‘world out there’ to which those endeavors are meant to provide access. As discussed earlier, the ‘pioneers of understanding’ have alerted us to the fact that this outlook engenders a preoccupation with raising bulwarks against skepticism (Zagzebski), leading to an underappreciation of both the cognitive achievements of science (Elgin), as well as of the problem of epistemic value (Kvanvig). The hope was that, through a thorough reconsideration of epistemology’s aims and methods, the discipline could be steered into more fruitful territory.

As regards the present investigation, the two concerns just discussed require that we address two fundamental issues of methodology: we need to reassess what our target of analysis is to be, and what purpose the analysis is supposed to serve.<sup>33</sup> With respect to the first issue, one suggestion that has been adopted by some is to shift the attention either towards minimally, or towards maximally demanding notions of understanding. Doing so enables one to account for degrees of understanding in terms of improvements upon some threshold level (*minimal understanding*), or as approximations to an optimum (*maximal understanding*), respectively.<sup>34</sup> Such proposals, which recast the target of inquiry, in principle leave intact the theoretical objective of providing a definition. As to the second issue, a possible way forward

---

<sup>33</sup> The second part of this question echoes the call for transparency implicit in Baumberger’s (2019, p. 370) complaint that “it is unclear how the dispute about the correct account of understanding can be resolved since it is unclear what exactly such an account aims at and thus what its criteria of adequacy are”.

<sup>34</sup> For examples of such views, see Grimm (2017) and Kelp (2015), respectively.

## 1 The study of understanding

is to sacrifice the philosophical aim of providing normatively binding criteria in order to make room for the gradual and context-dependent character of understanding. One way of doing this is to identify so-called *evaluative dimensions* along which understanding can vary in strength. According to this approach, which has been defended by Baumberger (2019), evaluative dimensions should be thought of as values rather than as criteria, which can be weighed differently across different contexts and may allow for trade-offs.<sup>35</sup>

Let me now elaborate on the methodology of my own project, starting with the specification of the target of analysis. First of all, as noted earlier in this chapter, the present study revolves around two concepts: symbolic and factual understanding. I aim to explicate both these concepts, in the Carnapian sense I shall elucidate below. Importantly, the two components of this intellectual exercise are linked: the explication of factual understanding is meant to demonstrate the fruitfulness of my explication of symbolic understanding. That is, I explicate factual understanding on the basis of my preferred account of symbolic understanding, and argue that there are significant theoretical benefits to the account of factual understanding so constructed. The idea is that these benefits indirectly confer justificatory support upon my account of symbolic understanding, given that my explication of factual understanding hinges upon that account. Hence, rather than constituting parallel investigations, the explications of symbolic and factual understanding are asymmetrically connected: my explication of symbolic understanding forms the basis for explicating factual understanding, and the fruits of that latter endeavor in turn serve to legitimize the results of the former. This dialectical approach is reflected in the structure of my dissertation: symbolic understanding is targeted in chapter 3; chapter 4 is devoted to factual understanding.

There is a strategic, as well as theoretical reason why I start with the concept of symbolic understanding, and then work my way towards an account of factual understanding. The strategic reason is that symbolic understanding has thus far received comparatively scant interests from epistemologists, and that this very lack of attention renders the concept suitable as a starting point for an inquiry with revisionistic aspirations. The idea behind this is that if we want to avoid replicating the problems associated with the conservative approach to the study of understanding, it is useful to select as our starting point that variety of understanding the study of which has been least influenced by the assumptions and methods of post-Gettier epistemology.<sup>36</sup> I thus take the liberty to exploit a fortunate implication of an

---

<sup>35</sup> Baumberger mentions three such evaluative dimensions: factivity, grasping and justification. As he explains (Baumberger 2019, p. 376), the way these dimensions operate in assessments of understanding can be likened to the way in which Thomas Kuhn (1977a) has suggested that criteria such as simplicity, accuracy and fruitfulness guide scientific theory choice.

<sup>36</sup> The relative neglect of symbolic understanding is likely due to a presumption on the part of epistemologists that developing a theory of that concept is not really a task for epistemology, but rather one for a linguistics, semiotics or cognitive science. Baumberger

otherwise unfortunate myopic outlook. However, my dialectical approach is not solely grounded in opportunism. For starting with symbolic understanding also makes sense from a theoretical viewpoint. After all, in light of my objective to show that the concept of factual understanding can be explicated on the basis of an explication of symbolic understanding, it is reasonable to say that the study of symbolic understanding should precede the study of factual understanding.

A further comment on the specification of my target of analysis is in order. One could reasonably ask, after all, what sort of attributions of symbolic and factual understanding I take as primary. Should we focus, for instance, on what it means to have symbolic or factual understanding simpliciter, or had we better, as some have proposed, tackle minimal or maximal conceptions of these concepts? I follow Kelp (2017) in designating *comparative assessments* of symbolic and factual understanding as the core units of analysis. This choice is based on the consideration that judgments to the effect that someone understands something better than someone else, tend to be easier to make and interpret than judgments to the effect that someone understands something outright, or to a certain non-comparative degree (idem, p. 266). This is because absolute assessments (whether outright or in terms of non-comparative degrees) are in fact implicit comparisons, which leave out the comparison class for want of determinacy or conversational relevance. Unlike assessments that come with explicit comparison classes, therefore, determining whether some absolute assessment of understanding is warranted requires making additional reasoned judgments about what the relevant comparison class is. The following analogy, due to Kelp (2017, p. 269), can serve to illustrate this:

For instance, the judgement that A is taller than B is arguably less complex than the judgement that A and B are both tall. Typically, all we need to get the former right is some way of knowing the height of A and B. As opposed to that, in case of the latter, we need to know not only the height of A and B but also various facts about the conversational context such as comparison class and standard of comparison.

In short, the decision to focus on symbolic and factual understanding, in conjunction with my choice to take comparative assessments as basic, entails that I will in this dissertation try to spell out what it means for someone to be a superior symbolic or factual understander, compared to someone else.

So much for the delineation of the target concepts. What will be the purpose of targeting them? It is worth pointing out that this question is not routinely posed in

---

(2014, p. 70), for instance, says that “as an epistemologist I am primarily concerned with factual understanding. Symbolic and, more specifically, linguistic understanding are the subjects of semiotics and the philosophy of language, respectively.” In a way, this attitude is revealing of the way in which the traditional knowledge-oriented paradigm continues to linger in the epistemologist’s professional self-image. For implicit in Baumberger’s words is the idea that the epistemologist’s core duties still lie exclusively with securing the link between our cognitive endeavors and the ever-elusive facts of the world.

## 1 The study of understanding

epistemological research that employs the method of conceptual analysis. Defining a concept in terms of necessary and sufficient conditions is often deemed to constitute a legitimate goal in itself, which need not serve any purpose other than establishing congruence between the verdicts of our intuition and the extension of the concept at issue. But such an approach, I have already indicated, is not to be replicated here. Instead of conceptual analysis, I will (once more) follow Baumberger (2019) in his suggestion to employ the method of *explication*, as understood along the lines proposed by Rudolf Carnap. According to Carnap (1950, p. 1), explication amounts to “the transformation of an inexact prescientific concept, the *explicandum*, into an exact concept, the *explicatum*”. What is required for an explicandum to be successfully explicated? According to Carnap (*idem*, p. 5), an explication needs to meet the following conditions of adequacy: “(1) similarity to the explicandum; (2) exactness; (3) fruitfulness [and] (4) simplicity.”<sup>37</sup> What is crucial about these criteria is that they allow for trade-offs: for instance, a decrease in similarity may be offset by an increase in exactness or fruitfulness. It is in this respect, also, that explication differs from conceptual analysis. Where the latter aspires to maximal preservation of our intuitive judgments, the former allows us to sometimes bracket those judgments if doing so contributes to achieving one or more of our other aims. For this reason, explication can be characterized as a kind of “voluntaristic conceptual engineering” (Leitgeb & Carus, 2021).

Together with my choice to let the study of symbolic understanding precede the study of factual understanding, the decision to employ the method of Carnapian explication is meant to effectuate a distancing from the post-Gettier epistemological tradition, with an eye to easing off the restrictive influence that tradition continues to wield over the study of understanding. In particular, the idea is that Carnapian explication might help us to conceptually engineer our way out of a paradigm that has thus far hindered us in realizing the prospects raised by the early proponents of the turn to understanding. This, however, is a very general goal that needs to be supplemented with a picture of how, concretely, this inquiry will seek to fulfil the criteria associated with successful explications. I will discuss this in the next section.

### 1.5 Specifying the aims of this project

Before my explicative endeavors can commence, we must have at least a rough idea of what it is that needs to be explicated. Carnap (1950, p. 4) has the following to say about this:

There is a temptation to think that, since the explicandum cannot be given in exact terms anyway, it does not matter much how we formulate the problem. But this would be quite wrong. On the contrary, since even in the best case we cannot reach full exactness, we must, in order to prevent the discussion of the problem from becoming

---

<sup>37</sup> There is considerable debate about how these criteria should be understood. In section 1.5, I will comment on how I interpret them.

entirely futile, do all we can to make at least practically clear what is meant as the explicandum.

What does ‘making practically clear’ consist in? According to the examples Carnap (idem, p. 4-5) provides, the prior clarification of the explicandum should be thought of mostly as a way of ridding ourselves of confusing ambiguities.<sup>38</sup> So, we may expect the prior clarification of an explicandum to take roughly the following form: ‘the explicandum is formulated as  $X$ , by which is meant  $Y$  rather than  $Z$ ,  $\phi$  rather than  $\psi$ ,  $P$  rather than  $P'$ , etc’. In the first part of this section, I will offer such prior clarifications for the pretheoretic notions we have of symbolic and factual understanding. In the second part, I shall comment on the way in which the intended explicata are meant to satisfy Carnap’s criteria for successful explications.

In the first section of this chapter, I indicated that symbolic understanding pertains to our understanding of languages and symbol systems, as well as to specific items within those languages and symbol systems. In this investigation, I confine myself to the understanding of specific linguistic and symbolic items. The account of symbolic understanding I develop is therefore not meant to bear directly on the understanding of languages and symbol systems. To be sure, my choice to concentrate exclusively on the understanding of individual linguistic and symbolic items does not mean that I believe that this kind of understanding can be had independently from having understanding of languages or symbol systems; such a claim would amount to a controversial denial of (even the more plausible versions of) semantic holism. Instead, my focus on the understanding of linguistic symbolic items only hinges on the trivial assumption that there is something for those items to be understood – period. Thus, there is something for this sentence to be understood, as well as for the words and phrases within the sentence. Whether or not these items can, in practice, be understood independently from the larger systems they are part of, and to what extent, is a separate question. Indeed, the view I develop in chapter 3 can be reconciled with different commitments regarding the subject of semantic holism, leaving open the possibility that individual linguistic and symbolic items can only be understood when one understands a language or a symbolic system.

A further point made in section 1 was that symbolic understanding intuitively has a behavioral as well as a cognitive (e.g. belief-based) component. Understanding a word, for instance, would seem to require an ability to apply it in suitable contexts,

---

<sup>38</sup> The relevant passage reads: “I might say, for example: “I mean by the explicandum ‘salt’, not its wide sense which it has in chemistry but its narrow sense in which it is used in the household language.” This explanation is not yet an explication; the latter may be given, for instance, by the compound expression ‘sodium chloride’ or the synonymous symbol ‘NaCl’ of the language of chemistry. (...) By explanations of this kind the reader may obtain step by step a clearer picture of what is intended to be included and what is intended to be excluded; thus he may reach an understanding of the meaning intended which is far from perfect theoretically but may be sufficient for the practical purposes of a discussion of possible explications.”

## 1 The study of understanding

and in the right ways, but at the same time it may be thought to demand the presence of appropriate cognitive states: one should be internally aware of the meaning of the word. Intuitively, then, there is an ‘active’ as well as a ‘passive’ ingredient to symbolic understanding. Heeding this intuitive picture, and factoring in the decision to take *better* symbolic understanding as the main unit of analysis, I shall now try to formulate the explicandum of symbolic understanding in clearer terms. For ease of reference, I refer to both linguistic and non-linguistic symbolic items by the general term *symbols*, and denote these using the letter  $\Phi$ .

### Explicandum of symbolic understanding

A subject  $S_1$  has a better understanding of some symbol  $\Phi$  compared to subject  $S_2$  if and only if compared to  $S_2$ ,  $S_1$  (i) exhibits, or has exhibited, superior behavior regarding  $\Phi$ , and/or (ii) has, or has had, superior cognitive states concerning  $\Phi$ .

Two disclaimers are in order regarding this formulation of the explicandum. First of all, I have had to resort to the inelegant ‘and/or’ construction, because of the fact that better symbolic understanding may sometimes be a function of both (i) and (ii), and sometimes only of either (i) or (ii). For instance, if neither  $S_1$  nor  $S_2$  is able to use  $\Phi$  in the right way, but  $S_1$  nonetheless has a better internal comprehension of  $\Phi$  than  $S_2$ ,  $S_1$  may possess a better understanding of  $\Phi$  by virtue of her internal comprehension alone. In other scenarios, when  $S_1$  as well as  $S_2$  possess at least some mastery of  $\Phi$  at both the active and passive levels, an attribution of better symbolic understanding to  $S_1$  may require that she outperforms  $S_2$  in terms of both (i) and (ii). Secondly, the word ‘superior’ is deliberately left unspecified. Nothing more concrete is needed here. For the purpose of clarifying the explicandum, there is no harm in assuming that there is such a thing as displaying comparatively superior behavior, or having superior cognitive states, without having a clear idea of what this involves.

Whether the above rendition of the explicandum indeed captures our pretheoretic notion of symbolic understanding can be tested through answering the following question: if (i) and/or (ii) are not met, should an agent always be denied symbolic understanding? My formulation of the explicandum answers in the affirmative. This, however, appears to be the wrong answer. After all, we need to take into account that intuitively, symbolic understanding pertains to the *possible* rather than to the *actual*. That is, attributions of symbolic understanding do not apply to one’s actual cognitive states or one’s actual behavior, but instead to one’s propensity to form the right cognitive states and to display the right behavior under the relevant circumstances. Or, to employ a distinction frequently used in the literature, it is not so much (past or present) *occurrent* behavior and cognition which are relevant to symbolic understanding, but rather *dispositional* behavior and cognition. This occasions the following revision to the initial formulation of the explicandum:

*Explicandum of symbolic understanding (improved)*

A subject  $S_1$  has a better understanding of some symbol  $\Phi$  compared to subject  $S_2$  if and only if compared to  $S_2$ ,  $S_1$  (i) has a disposition to exhibit superior behavior regarding  $\Phi$ , and/or (ii) has a disposition to form superior cognitive states concerning  $\Phi$ .

This new formulation of the explicandum says that dispositions (behavioral and/or cognitive) are both necessary and sufficient for symbolic understanding. Does this mean that occurrent behavior and cognition play no role in the explicandum of symbolic understanding? Yes, it does indeed: occurrences do not constitute symbolic understanding, but only ever provide evidence for its presence or absence. For if, in the relevant circumstances, a subject actually displays superior behavior and/or has superior cognitive states, this but confirms the hypothesis that she possesses dispositions to that effect. And conversely, if a subject fails to display superior behavior and/or have superior cognitive states under the relevant conditions, then this entails that she lacks the required dispositions. Either way, dispositions pre-empt occurrences.

Having clarified the explicandum of symbolic understanding, we can turn to the explicandum of factual understanding. Recall that factual understanding comprises that subcategory of objectual understanding which has as its grammatical object nouns or noun phrases that do not refer to symbols and symbol systems, but rather to things and phenomena ‘out there’.<sup>39</sup> Coining an umbrella term, let us call such items *factual targets*. It seems that unlike symbolic understanding, factual understanding is intuitively a uniformly cognitive affair. To understand the Suez crisis, or to understand the diffraction of light, appears to be a matter of believing, or otherwise cognitively apprehending, certain things (typically, but maybe not exclusively: explanations). Factual understanding lacks the distinctively behavioral component that features in symbolic understanding. Yet factual understanding is also similar to symbolic understanding in the sense that both revolve around dispositions. After all, in order to be credited with understanding the Suez crisis or the diffraction of light, one need not *here and now* have appropriate cognitive states regarding those subjects, but one must be disposed to form those states in certain circumstances. This insight, together with my earlier resolution to focus on comparative assessments of factual understanding, prompts me to formulate the explicandum of factual understanding in the following way, using the letter  $\Psi$  to denote factual targets:

---

<sup>39</sup> As said in section 1.1, I will return to the demarcation of factual understanding in chapter 4.



## 1 The study of understanding

### Explicandum of factual understanding

A subject  $S_1$  has a better understanding of some factual target  $\Psi$  compared to subject  $S_2$  if and only if compared to  $S_2$ ,  $S_1$  (ii) has a disposition to form superior cognitive states concerning  $\Psi$ .

Again, there is no need to have a clear idea of what it means to have superior cognitive states. I have merely endeavored to indicate what it is that needs to be explicated, with an eye to eliminating any confusing ambiguities.

With the explicanda clarified, I can elaborate on the main goal of this project: explicating the explicanda. In the previous section, I mentioned that for Carnap, successful explications require that explicata are sufficiently similar to their explicanda, as well as exact, fruitful and simple. With respect to the present investigation, I interpret these demands as follows:

- (1) The explicata of symbolic and factual understanding must be similar to their respective explicanda in the sense that the explicata's domains of application should have at least some overlap with the explicanda's domains of application.<sup>40</sup>
- (2) The explicata must be exact in the sense that the explicata should enable us to make more determinate and more fine-grained comparative assessments of symbolic and factual understanding relative to the explicanda.
- (3) There are two fruitfulness criteria pertaining to each of the explicata separately:
  - a. The explicatum of symbolic understanding must be fruitful in the sense that it should offer a means for formulating an explicatum of factual understanding.
  - b. The explicatum of factual understanding so formulated must be fruitful in the sense that it should enable us to provide answers to pertinent questions, or resolve pressing problems, relating to the study of understanding in epistemology.
- (4) The explicata must be simple in the sense that they should provide unified accounts of comparative assessments of symbolic and factual understanding across a range of contexts and disciplines.

---

<sup>40</sup> In *Logical Foundations of Probability* (1950, p. 7) we find the following, somewhat inconclusive remarks concerning the criterion of similarity: "in most cases in which the explicandum has so far been used, the explicatum can be used; however close similarity is not required, and considerable differences are permitted". The usual way of reading this is that dissimilarity is allowed to the extent that that enables us to better meet some of the other criteria (i.e. exactness, fruitfulness and simplicity). At least some measure of similarity must be ensured, lest an explication becomes liable to the objection that it merely changes the subject (Dutilh Novaes, 2020, p. 1016).

## 1.5 Specifying the aims of this project

Below, I provide cursory characterizations of the explicata, and comment on how these explicata are intended to satisfy the criteria of similarity, exactness, and simplicity, and how they meet their associated fruitfulness demands. Before I do so, however, I want to devote some attention to the main methodological principle by which I abide throughout this inquiry.

I have said that the standard, intuition-based method to epistemological theorizing is ill-suited for realizing the prospects that have motivated a number of scholars to propose a ‘turn to understanding’ in epistemology. My preferred alternative to the standard method is to forgo excessive reliance on our intuitions, and seek instead to account for the purpose(s) that ascriptions of symbolic and factual understanding serve to a community. Stated differently, the explicata should make sense of how such ascriptions contribute to realizing the goals that are associated with the use of the two concepts. My reason for adopting this methodological maxim is that it is more objective than its immediate contender, in the sense that claims about the goals of understanding-ascriptions are more readily verifiable (and falsifiable) than claims based on intuitive judgments. Whereas intuitive judgments are often difficult to contend with directly, claims about goals can in principle be scrutinized on the basis of intersubjectively available information about what those goals, in fact, are. The focus on goals instead of intuitions also implies that my inquiry has a significant descriptive aspect. But it is by no means exclusively descriptive. Although my explicata are meant to capture facts about certain practices (the descriptive aspect), they also have a clear normative import: the explicata are couched not in terms of dispositions individuals actually have (or tend to have) when they are credited with understanding by their peers, but rather in terms of dispositions individuals *should* have in order to be credited with symbolic or factual understanding. My objective is thus to offer normative accounts of symbolic and factual understanding, but normative accounts that are calibrated with the goals the use of those concepts can be seen to serve within, and for, a community.

That being said, I should note that I will not go so far as to provide empirical data in support of my claims about the goals of understanding-ascriptions. Rather, I formulate hypotheses about what these goals are based on some highly general observations. This ‘speculatively empirical’ method, while decidedly non-standard, has its precedents. Edward Craig (1990, p. 2), for instance, who is concerned with knowledge rather than understanding, adopts a policy similar to mine: “We take some prima facie plausible hypothesis about what the concept of knowledge does for us, what its role in our life might be, and then ask what a concept having that role would be like, what conditions would govern its application.” In turn, Craig’s approach to analyzing the concept of knowledge bears a resemblance to Wittgenstein’s *modus operandi* in his *Philosophical Investigations*. Wittgenstein’s aim is to show that in order to form an accurate picture of what it means to understand something, we should conduct an investigation into the *grammar* of the word ‘understanding’ (and its cognates): a scenario-based study of how that word actually

## 1 The study of understanding

operates in our language. As McGinn (2013, p. 113-114) explains in her guidebook to the *Investigations*:

Wittgenstein is concerned with a grammatical investigation of how our concept of understanding actually functions, which is aimed at drawing our attention to the kind of concept (...) that the concept (...) of understanding is; his remarks are purely descriptive of how we operate with these words, how we are taught to use them, and so on. The whole purpose of his description, like the purpose of his grammatical enquiries in general, is to reveal that the picture we're inclined to construct in response to the question, 'What is understanding?' is empty, and that everything that we need to remove our puzzlement lies already open to view in the way we operate with the expressions 'Now I understand', 'Now I can go on', and so on.

Although I do not necessarily endorse the conclusions that may be drawn from Wittgenstein's seemingly scattered observations about the grammar of the word 'understanding', I do concur with the general spirit of his philosophical method – a spirit that also permeates Craig's work: namely the idea that we should strive to account for the function of a concept within a practice, by tailoring our analysis of that concept (or, as is perhaps more accurate in Wittgenstein's case: the elimination of a faulty analysis) to the concept's function.

What are the functions of the concepts of symbolic and factual understanding? Before answering this question, I address a more fundamental one: what is the overarching function of understanding-ascriptions? In order to answer *this* question, I make two basic assumptions. First, I assume that an ascription of understanding amounts to evaluating something, or someone, in a positive way. Understanding, in other words, is a success term. Secondly, I assume that the aim of expressing a positive evaluation is to reinforce whatever it is that is being evaluated positively. That is, if we express our approval of something, I take it that we do so in order to increase the likelihood that what we approve of continues to exist, or occurs again. Based on these assumptions, I submit that we had best think of the general function of understanding-ascriptions (symbolic and factual alike) in the following way:

### *The general function of understanding-ascriptions*

An ascription of understanding to an individual serves to encourage that individual to sustain the possession of a certain disposition.

In short, to ascribe understanding to someone is to motivate her to continue having a particular disposition she has (be it a behavioral or a cognitive disposition). This means that understanding-ascriptions are implicit requests: to ascribe understanding is to say something like 'Please continue behaving, or believing, as you are disposed to do'. Conversely, to ascribe to someone a lack of understanding, is to discourage the sustainment of a certain disposition, and/or to encourage the formation of a new one. Regardless of how we conceive of the specific functions of symbolic and factual

understanding, this is what I take understanding-ascriptions to be: they are instruments for encouraging the possession of certain dispositions.

What, then, about the specific functions of ascriptions of symbolic and factual understanding? Here, I shall simply state what I think those functions are, postponing the provision of arguments in support of my claims as regards these functions to chapters 3 and 4, respectively. First, symbolic understanding:

*The function of ascriptions of symbolic understanding*

An ascription of symbolic understanding to an individual serves to encourage that individual to sustain the possession of a disposition that contributes to maintaining a stable practice of symbol-usage within a community.

Next, factual understanding:

*The function of ascriptions of factual understanding*

An ascription of factual understanding to an individual serves to encourage that individual to sustain the possession of a disposition that contributes to optimizing a community's attunement to its environment.

Both these conceptualizations of the functions of symbolic and factual understanding implicate behaviorist accounts of these concepts. That is, they entail that ascriptions of symbolic and factual understanding apply to behavioral (and not to cognitive) dispositions. I will discuss this implication in chapters 3 and 4. For now, I ask the reader to bear with me, as I use the remainder of this section to outline the essentials of the explicata I intend to formulate in this dissertation.

Very roughly, the anticipated end result of my explication of better symbolic understanding is the following:

*Explicatum of symbolic understanding (rough version)*

An agent *A* has a better understanding of some symbol  $\Phi$  than another agent *B* just in case *A* has a comparatively stronger disposition to competently use  $\Phi$  than *B*.

Without trying to spell out the meaning of this phrase in full, let me briefly elucidate four of its core concepts. The notion of an *agent*, first of all, is not meant to be restricted to human beings, nor even to beings that are capable of acting with (conscious) intent. Rather, an agent, as I understand it, is an entity – typically an organism – capable of purposeful action. In section 2.1, I explain in more detail what purposeful action, and therewith agency, consists in. Secondly, the notion of a *symbol*. In everyday talk of symbols, we usually think of items that denote their targets according to codified rules. Paradigmatic examples are the symbolisms of formal logic and the pictorial symbols used in cartography. As already hinted at in

## 1 The study of understanding

section 1.1, however, the concept of symbolic understanding implicates a much wider conception of symbolicity. According to this conception, any way of referring to something on the basis of explicit or implicit conventions – be it through ordinary denotation, mapping or modelling – involves symbols. Words, sentences, diagrams, gestures and physical models are, on this picture, all different varieties of symbols. In section 2.5, I formulate a definition of this wide-scoped concept of symbols. Thirdly, I define *usage* as the selection of means for purposes. Such selection may involve conscious intent on the part of the agent, but this need not be the case. In section 3.1, I present my definition in more detail and discuss its implications. *Competence*, finally, will be understood as propriety in symbol-usage which is due to an appropriate cause. This idea of competence, which draws on Ernest Sosa's theory of competent performance, will be touched upon in sections 3.3 and 3.4.

The account just sketched aims to fulfil Carnap's four criteria as follows. Let us consider similarity first. On the one hand, I strive to preserve an important aspect of the explicandum that was discussed previously: the idea that symbolic understanding is a matter of dispositions. This aspect will surface in section 3.1, where I stipulate that symbol-usage consists in a disposition to select a symbol as a means to achieving a purpose. On the other hand, the above description of the explicatum also implies a significant departure from the explicandum. For the fact that the notion of use pertains solely to behavior is at odds with the fact that our intuitive conception of symbolic understanding hinges upon behavior *and* cognition. In section 3.2, I argue that the cognitive aspect of symbolic understanding can be accounted for by seeing that internal comprehension is instrumental to acquiring dispositions for competent symbol-usage. On this view, internal comprehension is a contributing factor to, not a constituent of, symbolic understanding.

What is thus lost on the side of similarity to explicandum, however, is meant to be compensated for by the explicatum's exactness and fruitfulness. With respect to exactness, I show that the explicatum allows us to spell out in more detail the dimensions along which one's symbolic understanding can vary in comparative strength. More concretely, I argue that one's comparative degree of symbolic understanding is a function both of the extent to which one's disposition to use a symbol covers that symbol's entire range of application (*breadth*), and of how economical one's efforts in forming that disposition were (*depth*). Both these dimensions of evaluation will be given fairly detailed characterizations (see sections 3.6 and 3.7), enabling us to make assessments of symbolic understanding that are more determinate and fine-grained relative to the corresponding capabilities afforded by the explicandum.

The requirement of fruitfulness will be met in an indirect way, in the sense that the fruitfulness of the explicatum of symbolic understanding is meant to consist in the way in which it enables us to formulate an explicatum of factual understanding which has various theoretical benefits. The fruitfulness of the explicatum of symbolic understanding therefore ultimately derives from the benefits that are

associated with the explicatum of factual understanding. This does not mean, however, that my explication of symbolic understanding will be tailored to the goal of explicating factual understanding *ab initio*. Rather, my explicatory efforts will be geared towards the goal of accounting for the function of ascriptions of symbolic understanding. In my attempt to formulate the explicatum, I am thus guided solely by the methodological principle described previously, not by the fruitfulness criterion. That latter criterion does not operate as a practical guideline by which I abide in formulating the explicatum, but only as a standard in terms of which the result of my explication is to be assessed.

The criterion of simplicity, while arguably of secondary interest to Carnap, plays an important role in this project as well. However, my interpretation of this requirement is rather different from Carnap's. As Brun (2016, p. 1224) explains, Carnap seems to have thought of simplicity as a property of the syntactic form of a definition, rather than as a condition of ontological or conceptual parsimony. For the purpose of explicating symbolic understanding, however, I opt for the latter interpretation. While thereby admittedly deviating from Carnap's original method, this flexible interpretative choice allows me to incorporate as a condition of adequacy something which is often deemed to be an important theoretical value. The way I shall seek to satisfy this requirement is by offering an account of symbolic understanding that revolves around a very broad conception of symbols. In this way, the account unifies what might otherwise be considered distinct epistemic achievements. The proposed explicatum of symbolic understanding, while arguably lacking in simplicity as far as its syntactic form is concerned, is thus simple in the sense that it makes for a unified theory of what it is to understand various types of symbols.

Let me now turn to introducing the explicatum of factual understanding:

*Explicatum of factual understanding (rough version)*

An agent *A* has a better understanding of some factual target  $\Psi$  than another agent *B* just in case *A* has a comparatively stronger disposition than *B* to competently use symbols for the purpose of informing others about their environment.

Some of this explicatum's core concepts also figured in the explicatum of symbolic understanding, and require no further preliminary elucidation here. As regards those concepts that are new, a few comments are in order. First, the notion of a *purpose*, which is in fact central (albeit implicitly) also to the explicatum of symbolic understanding, will be defined in the context of my discussion of the agency in section 2.1. What is perhaps useful to note at this early stage is that a purpose, on my view, should be understood in a broadly Aristotelian way, according to which purposes are not psychological states (desires, intentions), but rather types of outcome. The second concept of importance, that of *informing others*, is understood

in a thoroughly pragmatist sense: to inform others is to enable them, by means of the competent use of symbols, to achieve certain purposes. More on this topic in chapter 4, in particular sections 4.2 and 4.3.

How is the explicatum of factual understanding meant to satisfy Carnap's desiderata? First, as with the explicatum of symbolic understanding, some measure of similarity to the explicandum is preserved by retaining the intuitive idea that factual understanding is a matter of dispositions. Admittedly, however, much of the intuitive picture is lost by virtue of the fact that factual understanding is explicated in terms of behavioral, rather than in terms of cognitive dispositions. The needed compensation of this deficit is meant to be procured by the explicatum's fruitfulness. As with the explication of symbolic understanding, the aforementioned methodological principle, which tells us to prioritize accounting for the function of ascriptions of factual understanding, will guide my attempt to formulate the explicatum. The theoretical benefits of the explicatum so formulated, reside in the explicatum's utility for resolving a number of important issues concerning understanding. In section 4.5, the following issues will be touched upon:

1. *The relation between symbolic and factual understanding.* I show that the explicatum of factual understanding, in conjunction with the explicatum of symbolic understanding, enables us to explain how these two epistemological categories are connected – a connection that would be difficult to account for in terms of the explicanda of both concepts.
2. *The epistemic acceptability of idealization.* I show that the explicatum of factual understanding allows us to accommodate the epistemic value of representational devices that deviate from the truth in some way or other, in a way that nonetheless makes sense of why those devices are, at the same time, also epistemically defective.
3. *The relation between objectual and explanatory understanding.* I show how the general methodology behind my explication of factual understanding can help to redirect counterproductive discussions over the purported primacy of objectual over explanatory understanding (or vice versa).
4. *Knowledge, understanding, and scientific progress.* I show that the proposed explicatum of factual understanding constitutes a welcome alternative to the traditional, truth-centered conception of knowledge, in the sense that it can more readily offer a story as to how our epistemic standings have improved, and perhaps also deteriorated, over time.
5. *The epistemic value problem.* I show that the explicatum of factual understanding provides a straightforward answer to the problem of epistemic value: having a better factual understanding implies being of more use to a community. In addition, I speculate as to how the superior

instrumental value of better factual understanding may ultimately derive from how factual understanding contributes to things of intrinsic value.

In a slight breach of Carnap's dictates, the criterion of exactness will play only a subsidiary role in my explication of factual understanding. Although the definition I formulate in section 4.2 goes some way towards improving upon the explicandum's (lack of) exactness, I do not spell out in detail the various dimensions along which one's factual understanding can vary in comparative strength. However, I do provide some suggestions as to how we could go about doing this. The simplicity criterion, finally, is heeded in much the same way that it is heeded in the context of explicating symbolic understanding: whilst the syntactic form of my definition is admittedly somewhat complex, the explicatum of factual understanding is simple in the sense that it unifies what might otherwise be thought of as distinct epistemic states or achievements (i.e. understanding phenomena, historical events, persons, subject matters, etc.).

This section should have enabled the reader to anticipate most of what follows in the rest of this dissertation. In closing, let me highlight the section's main elements once more. First, in line with Carnap's recommendations, I clarified the explicanda of the two target concepts of this investigation: symbolic and factual understanding. Then, I elaborated on the general methodological principle that governs my explication of both these concepts, namely the goal of tailoring the explicata to the functions the concepts can be seen to have in a community. I subsequently concentrated on how my explicata are meant to fulfil Carnap's demands. As part of this, I explained that the fruitfulness of the explicatum of symbolic understanding resides in its being instrumental to formulating a fruitful explicatum of factual understanding. In turn, the benefits of this latter explicatum were said to reside in its usefulness for solving a number of issues relating to the study of understanding.

## 1.6 Conclusion

The purpose of this prefatory chapter has been to demarcate the research topic of this dissertation, to situate my project within the contemporary epistemological literature, and to introduce its aims and method. As I explained, my research goal will consist in explicating the concepts of symbolic and factual understanding. How do we proceed from here? In the next chapter, I will attempt to define what a symbol is. I do so by drawing on the work of the American pragmatist Charles Sanders Peirce. Peirce's so-called 'semiotic', I shall argue, provides us with the resources for constructing a non-trivial account of symbolicity that captures all the various types of items falling within the purview of symbolic understanding.





## Chapter 2: Agents, signs, and symbols

In the previous chapter, I explained how this thesis connects to, and deviates from, work on understanding in the contemporary epistemology and philosophy of science literature. Furthermore, I indicated what sort of philosophical task I set myself to accomplish. The target concepts of this investigation will be symbolic and factual understanding, or, more precisely, *better* symbolic and factual understanding. Through utilizing Carnap's method explication, I will try to transform our pretheoretic conceptions of better symbolic and factual understanding (the explicanda) into more exact and fruitful explicata. As indicated in section 1.5, I shall explicate better symbolic understanding in terms of the comparatively superior usage of symbols by an agent. As a first step to working out this explicatum in full, the current chapter will attempt to spell out in detail, first of all, the notion of agency, and secondly, the concept of a symbol. The reason why I have chosen to execute this dual task within a single chapter, is that the two concepts can be clarified on the basis of a common theoretical framework. This framework is Charles Sanders Peirce's theory of signs, as interpreted and reconstructed by Peirce scholar Thomas Short.

**Section 1** argues that for the purposes of this project we can define an agent as an organism that engages in purposive behavior. Through reconstructing arguments due to Short, I show that purposiveness can be understood as a property which is attributed to a particular course of behavior based on there being some actual or modal pattern of behavior-adjustment of which that course of behavior forms part. Having addressed the concept of agency, I will turn to the second topic of this chapter: symbols. **Section 2** starts off from the assumption that symbols, however we define them, are things that 'stand for' other things. But what does this elusive relation consist in? Relying on Short's interpretative efforts, I suggest that Peirce's sign theory provides us with a suitably general and substantive account of signification. In **section 3**, I shall discuss two definitions that are central to Short's reconstruction of Peirce's concept of the sign. In **section 4**, I first attempt to show that if the use of symbols is to be something which can be governed by criteria of competence, then symbols must be types. I will then introduce two new notions, that of the type-sign and the token-sign, and try to characterize these on the basis of Short's definitions. Finally, **section 5** will, again on the basis of Short's work, provide definitions of symbols and their so-called replicas in terms of the previously formulated definitions of type- and token-signs. As a disclaimer, I want to make clear at the outset that while this chapter draws heavily on Peirce's sign theory via Short's interpretation of it, I do not claim to be providing an accurate representation of that theory. My goal remains to make a contribution to systematic epistemology, and it is to that end that I will employ Short's account of Peirce's theory of signs.

### 2.1 Agency

This chapter is largely devoted to matters belonging to the discipline of semiotics. The nature of the account I wish to defend, however, requires making a preparatory detour into adjoining territory. This concerns the meaning of the term ‘agent’, as it occurs in the explicatum of symbolic understanding. In contemporary epistemology, talk of agents, like talk of subjects, is most often meant to be about you and me, and about people sufficiently like us. In other words, the meaning of the term agent is often tacitly understood to be ‘ordinary human being’. To be sure, such implicit reliance on an anthropocentric understanding of agency may be perfectly benign. One is free to define one’s basic terms as one sees fit, after all. But in the present context, we should prefer a definition of agency that is in principle neutral with respect to the kinds of entities or organisms to which it applies. For a significant downside of relying on the anthropocentric conception is that it restricts the range of application of the explicatum a priori. It thus implicitly posits the existence of an epistemic divide between humans and non-humans. Now, replacing the anthropocentric conception of agency by one that has application beyond humans, does not imply that the scope of symbolic understanding is automatically widened accordingly. A more inclusive conception of agency opens up the way for a correspondingly inclusive conception of symbolic understanding, but the latter must ultimately be defended on grounds independent from the former. After all, even if agency is attributable to both humans and non-humans, there might still be sound arguments for claiming that symbolic understanding requires a particular sort of competence which only humans are capable of possessing. In order to at least leave room for settling this issue on the basis of considerations other than a priori distinctions, we had best adopt a non-anthropocentric concept of agency.<sup>41</sup>

In what follows, I will try to define the concept of agency in such a way that we avoid prematurely accruing substantial commitments regarding the nature of symbolic understanding through employing a restricted notion of agency. The current section therefore has a rather modest goal. One may be left wondering, however, why that goal should be pursued here, in the first section of a chapter on symbols. The main reason for this stems from the nature of the account of symbols which I present in this chapter. That account, which is rooted in Peirce’s theory of signs, takes symbols to be relative to purpose-directed responses. And the notion of a purpose-directed response presupposes some idea of what purpose-directedness as such consists in. Whilst the account can in principle be understood without a fully

---

<sup>41</sup> Of course, trivially, the *term* ‘symbolic understanding’ is not applied outside of communities of human beings (or even outside of English-speaking communities). What I mean, however, is that the *concept* of symbolic understanding, whilst here grammatically delineated on the basis of the English language, may in principle have application beyond humans. This is the case when a community of non-human agents uses some evaluative device which is delineated analogously and which functions in all the relevant ways like our concept of symbolic understanding.

specified concept of agency in mind, both the account's plausibility and the intelligibility of its implications are enhanced significantly if such a concept is settled on beforehand. So, in a way, the goal of this section is twofold: its primary aim is to formulate the first component of the explicatum of symbolic understanding, but in doing so it also prepares the ground for the theory of symbols that I develop in the ensuing sections.

One of the central questions of action theory has been aptly formulated by Harry Frankfurt (1978): "The problem of action is to explicate the contrast between what an agent does and what merely happens to him, or between bodily movements that he makes and those that occur without his making them." Following the publication of Elizabeth Anscombe's seminal work *Intention* (1957), much subsequent discussion in the philosophy of action has revolved around intentional action, rather than around action simpliciter. Importantly, the notion of intentional action invites what Stout (2005, p. 6) calls an 'inward-looking' perspective on action: a perspective that aims to account for the difference between action and non-action in terms of some mental difference, and that endeavors to explain the role mental states play in bringing about actions. While such issues need not be unsolvable, they have proven to be notoriously hard to tackle. How do we individuate, let alone locate, mental states? And how could something mental cause, or give rise to, something non-mental? Apart from the fact that I do not wish to get involved in such questions, taking intentional action to be the hallmark of agency would thwart our purpose of adopting a non-anthropocentric account of agency. Intending, after all, is by all appearances a mostly human affair. Therefore, my target concept in this section will be situated one level 'below' that of intentional action. That is, I will try to say more about agency in its sense as 'purposeful, or goal-directed activity' – whether intentional or not. As with almost everything else in this chapter, my main intellectual source will be the work of Peirce scholar Thomas Short. Short's thoughts on agency venture beyond those of his main research subject, to the point where we might say that his work on this topic is as much systematic as it is historical. For this reason, and as a matter of convenience, I will in this section leave out reference to those Peircean roots from which Short's ideas might have originated.

Let us start off with some of Short's (2007, p. 92) basic observations regarding the meaning of the word 'purpose':

A purpose is always general, never a particular. A purpose has to be general, since, when it is a purpose, it is not yet attained: it is not yet actual, and only the actual or what was actual is particular. What is yet to be done, no matter how exactly we specify it, leaves room for infinite variations, even if minute and unimportant. Of course, normally, one seeks an outcome that will be a particular, but what one seeks is a particular – any particular – of a given type. It is the type alone that can be specified; before the type is achieved, there is no particular of that type that could be specified.

## 2 Agents, signs, and symbols

A purpose, in short, is a type of outcome. But it is of course not just any type of outcome. The creation of a black hole may be a type of outcome of the death of a star, but it evidently is not a purpose of the star's dying. What distinguishes purposes from other types of outcome? The answer cannot be that a purpose is a type of outcome which someone or something intends or desires to achieve. For we speak freely of organisms pursuing purposes whilst refraining from ascribing intentions or desires to those organisms. A silver eel making its way to the Sargasso sea clearly does so for the purpose of reaching a place suitable for reproduction, but we may suppose that it is mere instinct, and not conscious intent, which drives this behavior. So something can be a purpose without it being intentionally strived for.

Short (idem, p. 110) has suggested a different, and to my mind more plausible account of the nature of purposes. According to this account, roughly, a purpose is a type of outcome which someone or something tends (not: *intends*) to bring about across a variety of circumstances, by virtue of a disposition to flexibly modify its behavior under certain conditions. More succinctly put, purposes exist where flexibility in behavior leads to stability in outcomes. This idea is best illustrated by means of an example:

[I]f you place an obstacle across the path of a line of ants hurrying back and forth between their hill and some source of food, you will find that, after some random explorations along the edges of the obstacle, the ants will take a new path (if one still exists) between home and food. (Short, 1981a, p. 370)

Faced with a novel obstacle, a group of ants in search of food will tend to bring about the same type of outcome it brought about before the obstacle was put in place. We thus have a remarkable consistency in type of outcome: an alteration of the environment does not change the type of the event that follows, compared to the situation prior to the alteration. Why is this remarkable? Because ordinarily, alterations in the environment *do* tend to give rise to changes in outcome. High temperatures cause water to evaporate quickly; still higher temperatures cause it to evaporate even more quickly. Water does not have the ability to maintain a stable rate of evaporation through randomly varying its behavior in response to changed circumstances. Ants, like so many other organisms, are special in this regard: in at least certain contexts, they are capable of maintaining stable outcomes despite changes in their environment. And they are so capable because they can respond to environmental changes by varying their behavior, and by subsequently converging towards courses of behavior that lead to types of consequence that would have occurred absent the environmental changes.

To the picture just sketched, I need to add that a given type of outcome is a purpose of some entity or organism not only if that entity or organism *actually* maintains stability in type of outcome through variation of its behavior, but also if it would do so across various *possible* scenarios. To see this, think of the following

pair of examples: lifting a glass of water to one's mouth being followed by drinking the water, versus stumbling over a loose paving stone being followed by a painful encounter with the ground. The lifting of a glass is of a kind that is unlike the stumbling over a loose paving stone: it is, as Short (2007, p. 111) explains "subject to modification if it fails its purpose". If I pick up the glass too loosely, and subsequently drop it, chances are that I will try again applying a firmer grip (perhaps after first refilling it), in such a way that I am then more likely to have my thirst quenched. Stumbling is not similarly liable to correction: having avoided falling after stumbling over a paving stone, I do not walk back and adjust my stumbling in such a manner that falling is then more probable to occur. Failing to consume the contents of a glass invites renewed attempts in a way that failing to fall does not. Now, we usually succeed when we try to lift a glass and consume its contents. In ordinary life, patterns in types of outcome are often obscured by our fortunate disposition to do simple things right at the first attempt. Yet such patterns still exist in a modal sense: if failure were to occur, we would indeed modify our behavior in such a way that the consequences of our doings would continue to follow a particular trend.

As will have become clear, Short's account of purposes makes no reference to mental states. For Short, types of outcome are not purposes depending on whether they are desired or intended, but rather depending on whether an organism or entity's behavioral flexibility makes (or would make) those types of outcome occur consistently across a range of different (counterfactual) circumstances. Because it avoids reference to the mental, this notion of purpose has application beyond humans. Still, we should be wary of the range of application becoming too wide. Take the example of a river flowing out to sea. One could say that flowing out to sea is a type of outcome of the behavior of a river. For certain kinds of environmental change, that type of outcome tends to persist: if one places a rock of some size and weight somewhere in the river, for instance, the water will likely 'find' a way around it and continue on its trajectory towards the sea. Does this render 'flowing out to sea' a purpose of a river? As far as our ordinary usage of the word 'purpose' is concerned, it better had not. To exclude such examples as the one just sketched, we should further specify what we mean by 'a disposition to modify behavior'. As Short (*ibidem*) argues, in order to speak of a type of outcome being a purpose, the presence of a disposition for behavioral variation must be due to natural selection. That is, it should be the case that that disposition exists through having been selectively retained, by virtue of bestowing a reproductive advantage upon the organism possessing the disposition. Thus, while one may grant that a river is disposed to modify its behavior upon environmental changes, it would be wrong to say that a river is so disposed because that disposition was selected for. Consequently, flowing out to sea is not a purpose of a river.

## 2 Agents, signs, and symbols

Since only properties of biological organisms can be selected for, only biological organisms can – in principle – behave purposefully.<sup>42</sup> Their behavior is purposeful just in case the type of outcome which their behavior gives rise to occurs (or would occur) consistently by virtue of the organism's selected-for disposition to flexibly modify its behavior in response to changed circumstances. An agent, on this account, is an organism capable of purposeful behavior. Obviously, however, no organism will manifest agency in every aspect of its behavior. For example, humans may be agents when it comes to lifting glasses, but not when it comes to having goosebumps. Having goosebumps is a response to low temperature or emotional excitation which is not really liable to flexible modification. Thus whenever in the remainder of this dissertation someone or something is called an agent, this means that that someone or something is being considered *insofar* as it is capable of purposeful behavior.

Admittedly, the above definition is not completely watertight, since it may sometimes be up for debate whether some organism really has a particular disposition for flexible behavior-modification. Experience tells us that behavior which is in principle subject to modification is not *always* relevantly modified when disaster strikes. Within the set of possible scenarios associated with the act of lifting a glass of water to one's mouth, for instance, there may be instances in which one's clumsy dropping of the glass is followed by a walk in the park. However, this intuitively does not alter the fact that lifting the glass to one's mouth can be described as purposeful behavior. Being subject to modification is thus a matter of self-correction being the prevailing, rather than the exclusive response to failure. More specifically, we can say that some behavioral episode is subject to modification if, in a sufficient proportion of cases, failure to bring about some outcome is followed by an adjusted course of behavior from which that same outcome does follow, or is likelier to follow. When, exactly, there is such a sufficient proportion of cases is a matter not exhaustively determined by objective criteria. Indeed, to put it more strongly: judging whether some course of behavior is purposeful is a decidedly human affair. For any such judgment necessitates (implicit) mutual coordination both with other people's estimates of observed relative frequencies of self-corrective behavior and, more crucially, with their assessments of when a certain proportion of

---

<sup>42</sup> The criterion excludes as potential agents also those parts of organisms which may be said to exist for a purpose, but which cannot be said to behave purposefully independently of the organism they inhabit. The human eye, for instance, may be said to exist for particular purposes, but we should not say that it engages in purposive behavior. Of course, the typical behavior of the eye is a result of repeated natural selection over the course of millions of years. But it is evidently not in virtue of any reproductive advantage so bestowed upon the eye itself that such selection has occurred. Rather, it is the reproductive advantage bestowed upon the organism as a whole which explains why certain of its features have been selected for. See also Short (2007, p. 111) for a more detailed discussion of the distinction between existing and behaving for a purpose.

self-corrective behavior counts as sufficient.<sup>43</sup> Attributions of purposiveness are therefore prone to bear a hint of arbitrariness.

Let me now conclude this section by offering a summary of its contents. I started out by saying that we need an account of agency both in order to specify a component of our explicatum, and to set the stage for a discussion of Peirce's theory of signs in the remainder of this chapter. I pointed out that our theoretical goals are best served by departing from analytic philosophy's traditional focus on intentional action, and by concentrating instead upon the sense of agency as purposeful behavior. I then offered a reconstruction of Short's account of purposeful behavior, according to which some piece of behavior by an organism is purposeful just in case the organism is disposed, by virtue of properties selected for, to flexibly modify its behavior in such a manner that a certain type of outcome thereby occurs (or would occur) consistently.<sup>44</sup> While this definition might not allow us to adjudicate between all

---

<sup>43</sup> One could add to this that all this mutual coordination hinges on an ultimately contingent way of individuating events, outcomes, and patterns. We count the lifting of a bottle as the event of relevance, but why not the lifting of a glass plus the touching of one's nose? Attributions of corrigibility are thus dependent upon a certain way of carving up observational contents for which there in principle exist alternatives. But I take this to be a problem that applies to observation in general. Why do we perceive an item as an instance of type *X*, rather than as an instance of type *Y*? The question reminds one of canonical philosophical problems surrounding the general issue of underdetermination, such as Goodman's new riddle of induction, and Quine's indeterminacy of translation thesis.

<sup>44</sup> A brief disclaimer is in order regarding the way in which I have drawn upon Short's theory. In his 2007 book, the concept of a purpose is intimately tied to what Short calls *anisotropic explanation*, which is explanation of a phenomenon in terms of a type or type of effect exemplified by that phenomenon (Short, 2007, p. 130). According to Short, behavior qualifies as purposeful when an anisotropic explanation is needed to explain a behavioral pattern which cannot be accounted for in terms of a so-called *mechanistic explanation*, being an explanation of a particular outcome in terms of particular conditions (idem, p. 115). Purposes are thus 'real' to the extent that anisotropic explanations are called for to make sense of patterns in behavior. According to Short, the need to invoke anisotropic explanations arises when the degree of random behavioral variation manifested by an organism is such that the organism's tendency to reduce this variation to a type of behavior that yields a uniform outcome must be regarded as a form of selection. In other words, when a pattern of behavior is of such a nature that it resembles the kind of pattern effected by natural selection (i.e.: from random variation to uniformity), we must resort to anisotropic explanation – meaning that we need to explain the existence of the behavioral pattern in terms of its associated type of outcome.

By his own admission (Short, 2007, p. 141, and, more elaborately, Short, 1983, p. 318) Short's account of purposeful behavior goes beyond a theory such as Ernst Mayr's (1974), according to which certain behavior can be called 'teleonomic' when it involves the operation of a mechanism or 'programme' that was selected for certain outcomes it tends to produce. For whereas Mayr held that teleonomy involves seemingly goal-directed processes that are in fact fully mechanical, Short maintains that certain processes are indeed goal-directed in a more substantial sense. To an extent, Ruth Millikan's (2004, p. 3-14) discussion of the nature of purposes is similar to Short's treatment of the topic. Like Short, Millikan recognizes more than one level of selection, and accordingly seems to countenance talk of purposes in a non-Mayrian (that is, non-metaphorical) sense. Although Short may thus not be alone in his



cases, it is sufficiently discriminating to fulfil the dual function it is meant to be used for in this dissertation.

## 2.2 Peirce's concept of the sign

Earlier, I announced that the present chapter would attempt to draw more definite boundaries around the extension of the term 'symbol'. As indicated, this element of our explication of symbolic understanding should be understood in a very broad sense, according to which symbols include referential devices used in (natural) language, but also various forms of representation used in science. Before definite boundaries can be drawn, however, a rather fundamental issue needs to be addressed. For however broad or restricted we take the category of symbols to be, there is one feature that all symbols have in common: they somehow manage to refer to, stand for, or otherwise be associated with, something else. Let us call this elusive relation *signification*. What is the nature of signification, and under what conditions does it obtain? One way of avoiding excessive theory-building would be to designate the relation of signification as an undefined primitive. That may be an attractive strategy, given that it enables us to proceed swiftly. However, there is also a downside: with such a starting point, our account of symbolic understanding might come to rest on a rather arbitrary basis. To avoid this, I will walk a somewhat longer road. Over the course of two sections, I expound on the theory of signs developed by Peirce. The present section provides a first characterization of that theory, along with a diagram and a brief example, and argues that the theory is uniquely suitable for our purposes: it explains the nature of the signifying relation in terms of more primitive notions and supplies concrete criteria for when it obtains. In section 2.3, I specify the theory in more detail and illustrate it by means of an extended example.

For the medieval scholastics, signification was defined in terms of the simple formula *aliquid stat pro aliquo* ('something stands for something else'). This formula is neutral with respect to the origin of the 'standing for' relation, as well as to the nature of the relata. As a result, the formula is very permissive: something can function as a sign irrespective of the kinds of entities involved, and regardless of the way in which the sign-function is established. The medieval notion of signification should thus be thought of as covering not only ordinary *referential relations* (e.g. a proper name signifying a person or object), but also what we would nowadays regard as *evidential relations* (e.g. a medical symptom signifying a disease). The all-inclusive scholastic formula can be traced back to Augustine, who characterized the sign as "something that shows itself to the senses and something other than itself to

---

rehabilitation of teleology, I have refrained from explicitly grounding the proposed definition of agency in Short's background theory of anisotropic explanation, because of space limitations, but also in order to avoid rendering the definition any more controversial than it needs to be.

the mind" (*De Dialectica*, 1975, p. 86, as cited in Meier-Oeser, 2011).<sup>45</sup> Although Augustine did distinguish between those signs that exist independently of any purpose to signify (*signa naturalia*), and those that exist only in virtue of such a purpose (*signa data*) (Meier-Oeser, 2011), their common denominator of 'signhood' was of significant, if not superior, interest in itself.

The inclusive conception of the sign may strike one as outdated and therefore largely irrelevant for the contemporary systematic philosopher. After all, the advance of the exact sciences and the humanities has arguably rendered the medieval perspective obsolete: for quite some time now, we have had rather interesting things to say about (different forms of) evidence and reference considered as separate phenomena. A notable exception to this modern mode of thought is to be found in the work of Peirce. Intimately familiar with the writings of several medieval philosophers, Peirce developed an impressive theory of signs, or 'semeiotic' in his own words, that is similar in scope to Augustine's account.<sup>46</sup> Unfortunately, the prevailing attitude among 20<sup>th</sup>-century analytic philosophers towards Peircean sign theory seems to have been one of puzzlement-induced neglect. Only within the discipline of semiotics has Peirce's theory received ample attention, albeit not always of a sufficiently scrupulous sort.<sup>47</sup> Part of the explanation for both the

---

<sup>45</sup> Markus (1957, p. 71-72) notes that according to Augustine, signification is a relation between three relata, namely "the object or *significatum* for which the sign stands, the sign itself, and the subject to whom the sign stands for the object". As Markus (*idem*, p. 72) goes on to point out: "Augustine appears to be the first to have stressed this triadic nature of the relation of 'signifying': it had been noticed before that signs belong to the category of relation (...), but in all previous discussions the relation of sign to *significatum* is conceived of as a straightforward dyadic relation."

<sup>46</sup> With regard to Peirce's familiarity with medieval authors, Tiercelin (2006, p. 160) writes: "When one takes a look at the list of books by medieval logicians which are available in Harvard in 1880, one is impressed by the amount of rare books belonging to Peirce himself (295 volumes), books which he had acquired during his various stays in Europe and which were donated by his wife to the Johns Hopkins Library after his death: among the listed works and authors, one can find such names as: Boethius, Berangarius, Gilbert of Poitiers, John of Salisbury, Averroes, Peter of Spain, Alexander of Hales, Thomas Aquinas (7 books), Roger Bacon, Duns Scotus (five books), Ockham (5 books), Paul of Venice, etc."

<sup>47</sup> Semioticians study the general principles behind sign relationships and sign systems (*theoretical semiotics*), as well as the ways in which sign relationships operate in various domains of reality (*applied semiotics*). They tend to regard Peirce as one of the founding fathers of the discipline, together with the Swiss linguist Ferdinand de Saussure. Unfortunately, these two foundational influences have amalgamated into a supposedly omniable framework that combines the generality of Peirce's 'semeiotic' with the structuralism of De Saussure's 'semiology'. Armed with this framework, semioticians have in recent decades laid claim to intellectual territory not usually associated with sign systems or sign relations. This is reflected in the emergence of subdisciplines such as *neurosemiotics*, *phytosemiotics* and *sociosemiotics*. Together with a proliferation of elaborate taxonomies and obscure jargon, this development has contributed to the fact that semiotics currently enjoys a somewhat questionable reputation. The skeptical attitude towards semiotics finds an amusing expression in the remark, attributed to media theorist Paddy Whannel, that

apparent lack of interest among philosophers and the tendency for misinterpretation among semioticians, is to be found in the nature of Peirce's thought and work. He tended to continuously redevelop his ideas over the course of many unpublished and scattered manuscripts, which makes it difficult to reconstruct, let alone critically engage with, a definitive theory.

Since the commencement of the *Peirce Edition Project*, which strives to compile a comprehensive, chronological edition of Peirce's works, historical scholarship on Peirce's sign theory has notably matured. Throughout this chapter, I rely mostly on a recent interpretation defended by Short in his book *Peirce's Theory of Signs* (2007), from which I already cited some passages in the previous section.<sup>48</sup> Among other things, I shall in section 2.3 borrow from Short's interpretation two definitions that can serve to answer this section's opening question. The great advantage of these definitions is that they jointly make for a highly general and substantial picture of signification. As such, they constitute the distinctively Peircean groundwork for the account of symbolic understanding that I defend in this thesis. This is not to imply, however, that I thereby pledge allegiance to the metaphysical and phenomenological ideas that underpin Peirce's theory of signs. As is well known, Peirce's semeiotic is inextricably connected to his so-called 'phaneroscopy', which divides all of our experience into three categories: *firstness*, *secondness*, and *thirdness*. While fascinating both in its depth and originality, this aspect of Peirce's philosophy is also notoriously difficult to get to grips with. I will therefore not enter into that territory here. In further deviation from Peirce, sections 2.4 and 2.5 will feature a number definitions which have no obvious equivalents in Peirce's semeiotic, and which might not be altogether compatible with it. Hence, the connection between my own work and Peirce's philosophy is one of selective appropriation, rather than (purported) subsumption.

Although Peirce revised and augmented his theory several times, there is one assumption that remained consistently in place. According to this assumption, a sign is an element of a triadic relation which cannot be reduced to a combination of dyadic

---

"semiotics tells us things we already know in a language we will never understand" (Chandler, 2017, p. 8).

<sup>48</sup> In his 2007 book, Short starts out by saying that "Peirce's theory of signs, or semeiotic, misunderstood by so many, has gotten in amongst the wrong crowd. It has been taken up by an interdisciplinary army of 'semioticians' whose views and aims are antithetical to Peirce's own, and meanwhile it has been shunned by those philosophers who are working in Peirce's own spirit on the very problems to which his semeiotic was addressed" (Short, 2007, p. ix). From what follows, it becomes clear that the 'right crowd' Short is alluding to here, must be sought within the analytic tradition in 20<sup>th</sup>-century and contemporary philosophy. Part of Short's aim was to invite analytic philosophers to reconsider the merits of Peirce's sign theory through reconstructing that theory in more familiar terms, and by suggesting ways in which Peirce's work could contribute to the solution of well-known problems in philosophy of language, epistemology and philosophy of science.

or monadic relations.<sup>49</sup> Given that Peirce's work was continuously under construction, it is difficult to provide a direct citation that neatly captures the common denominator of all the different versions. I will therefore suggest the following as a preliminary approximation of Peirce's 'basic idea of the sign':

A *sign* is something which stands for an *object* to an *interpretant*.<sup>50</sup>

Obviously, this does not yet explicate the relation of reference, but I will turn to that shortly. I must first dispel potential confusion concerning the most esoteric component of this formula: the notion of an *interpretant*. On the one hand, the interpretant is not to be identified with an agent who does the interpreting. On the other hand, the concept of an interpretant should also be distinguished from our ordinary notion of an *interpretation*. Instead, an interpretant is (a feature of) a purpose-directed response exhibited by an agent, which manifests a particular interpretation of that which the agent responds to. 'Manifesting an interpretation' should be thought of as the actualization of a possibility – akin to the way in which an instance of some property manifests that property. For as Short (2007, p. 30) explains, interpretants more or less relate to interpretations as particulars relate to universals:

An interpretation is general, not particular: we may identify it, in Fregean style, as an equivalence class of interpretants or, more in Peirce's manner, as something embodied in or expressed by any of a number of interpretants (which then are, in that respect, equivalent). 'Interpretation' may of course refer to the process of interpreting, but 'an interpretation' and 'interpretations' can only refer to the products thereof. That product is general but borne by a particular, the interpretant.

Interpretants are often thoughts, but they can also be actions or feelings unaccompanied by thought or even conscious awareness.<sup>51</sup> The instinctive turning of one's gaze in response to someone else's pointing, for instance, is an example of

---

<sup>49</sup> Short (2007, p. 73) illustrates the difference between reducibly and irreducibly triadic relations by means of the following example: "[T]he triadic relation of benefitting reduces to a logical compound of two dyadic relations: x benefits y by z iff x did z & z benefits y. Giving, Peirce argued, is not similarly reducible, since x's laying y down and z's picking y up does not constitute giving." Signhood is like giving, and unlike benefitting, in this respect.

<sup>50</sup> When it comes to corroborating this rough formula with textual evidence, the following passage from Peirce (*CP* 2.228) is perhaps the most illuminating: "A sign, or *representamen*, is something which stands to somebody for something in some respect or capacity. It addresses somebody, that is, creates in the mind of that person an equivalent sign, or perhaps a more developed sign. That sign which it creates I call the *interpretant* of the first sign. The sign stands for something, its *object*. It stands for that object, not in all respects, but in reference to a sort of idea, which I have sometimes called the *ground* of the representamen."

<sup>51</sup> So, in keeping with what was said about purposefulness in section 2.1, interpretants are invariably purposeful, though not always produced with conscious intent.

## 2 Agents, signs, and symbols

an action-based interpretant in which the other's pointing is interpreted as being a directive or request of a particular sort (Short, 2007, p. 52). Seeing that the term interpretant may easily be interpreted mentalistically, I will from now on follow Short's convenient habit of using the more neutral term 'response' in referring to Peirce's interpretant. In order to preserve an important nuance contained in Peirce's own terminology, however, I should add to this that the term 'response' can henceforth refer either to a response in its entirety, or to some aspect of a response which is relevant to the sign relation at issue.

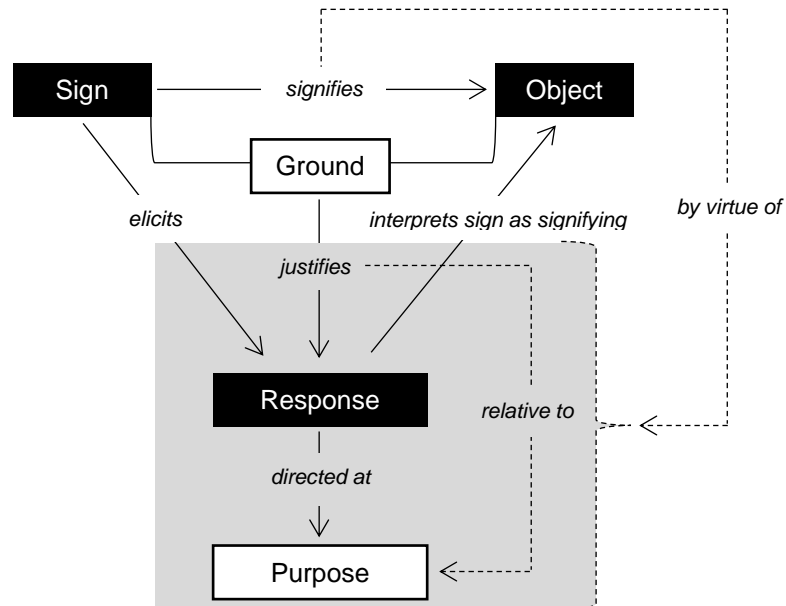
According to Peirce's account, it is relative to some thought, action, or feeling which interprets an item as a sign of something, that that item signifies what it is interpreted as signifying. But, as Short (1981b, p. 199) points out, "it is not in virtue of any possible interpretant that a sign signifies what it does. Rather, there must already be some relation between the sign and its object which determines or justifies a specific type of interpretant." Here we encounter an often overlooked element in Peirce's account of signification: the notion of a *ground*. A ground, according to Short's reading (2007, p. 162), is a relation obtaining between two items prior to any response, which justifies the interpretation of one of those items as being a sign of the other. The kind of justification at stake in such contexts is practical rather than epistemic: the interpretation of something as a sign of something else is justified to the extent that manifesting that interpretation in a response would constitute an appropriate means to achieving the response's purpose. Whether a particular interpreting response is so justified, will by and large depend on the probability that the response brings the agent at least somewhat closer to achieving the purpose for which that response was produced. I say 'by and large', because determining whether some response is appropriate in light of its purpose may further require that the response is weighed against possible alternatives which could also constitute means to achieving the agent's purpose. Such weighing will involve a comparison of the alternatives in terms of how likely the agent is to attain his purpose by means of each of them, and in terms of how efficient the various alternatives are as regards the expenditure of time and resources (idem, p. 155n1).<sup>52</sup>

We can sum up the foregoing as follows: according to (Short's interpretation of) Peirce, a sign signifies an object through being interpreted as doing so on the basis of a prior relation between sign and object, and by virtue of that interpretation being justified with respect to the purpose of the response in which the interpretation is manifested. In order to render this admittedly somewhat complicated idea more

---

<sup>52</sup> There is plenty of devilry in the detail here. As Short (2007, p. 154) explains: "There are all sorts of ways in which something may fit its purpose – being of the right shape or color or forcefulness, and so on. Fitness varies along another dimension as well: something may fulfil its purpose or be a means merely; and if a means, it may be sufficient or necessary or something that can be used for that purpose; it may be appropriate simply because nothing else is any better. The determination of what is appropriate depends on estimates of probabilities and cost/benefit analyses relating to entire sets of purposes."

readily comprehensible, we can try to visualize it by means of a diagram. Before doing so, however, it should be noted that any such attempt is bound to partly lose out on the intrinsically dynamic nature of Peirce's sign-concept. This dynamic aspect resides in the temporally extended character of signification: a sign is something which obtains relative to a response *which that sign elicits*, and which thus occurs only after the signifying item is encountered by the agent. Hence, one should keep in mind that the diagram below represents a process as much as a relation:



Perhaps as an additional aid to understanding the dynamic aspect of Peirce's sign theory, we can resort to an example. A helpful one from Short (*idem*, p. 156) involves a deer fleeing in response to the sound of a gunshot. On Short's rendering, the deer's response of running in a direction away from the sound, manifests an interpretation of that sound as a sign of a potentially life-threatening situation. This response is practically justified in case there is, let us say, a relation between gunshot-sounds and injury or death (e.g. a statistical correlation), which makes it so that running away is a prudent thing for the deer to do in case its purpose in running away is to avoid such unwelcome outcomes as injury or death. That is to say, if the deer's running away in response to the sound of the gunshot at least marginally increases the likelihood that the deer escapes injury or death, its act of running away is justified in light of that purpose. And if the deer's response is so justified, we can say the deer, in running away, justifiably interpreted the gunshot-sound as a sign of a potentially life-threatening situation.

A distinctive asset of Peirce's theory of signs is that it paints a highly general picture of what signs are, without incurring a significant loss at the level of the specificity of the criteria that determine when something is a sign. The account is general, because it recognizes any instance of interpretation as a potential case of

signification. But it is also remarkably concrete, since it explains how signhood depends on the concept of practical justification, for which we can adduce relatively straightforward conditions. This is not to say that it will always be easy to verify whether some interpreting response is justified in light of its purpose. But if we encounter any difficulties, these are more likely to stem from the verifying process being complicated (due to there being various alternative means to achieving the purpose, for instance), rather than from it being insufficiently specified. An added benefit of Peirce's account is that it enables us to both recognize the common denominator between relations of evidence and of reference, *and* acknowledge the difference between the two. However, as stated earlier, my interest in Peirce's theory lies in its capacity to provide us with a wide-scoped definition of symbols, and it is to that purpose that my use of the theory will be geared.

### 2.3 Subjective and objective signification

The previous section offered a summary version of Peirce's conception of signs. The current section will try to unpack that summary rendition, again on the basis of Short's reconstruction. Its main aim is to introduce a distinction between *subjective* and *objective signification*. Something is a subjective sign when it is *interpreted* as a sign by an agent, while something is an objective sign when it is *interpretable* as a sign. In what follows, I shall discuss two of Short's definitions that correspond to the notions of subjective and objective signification. These two definitions will constitute the theoretical backbone for the remainder of this chapter.<sup>53</sup>

The central idea of Peirce's theory of signs is that signification consists in a triadic relation between a sign, an object and an interpreting response. This basic idea was already clarified in the previous section. In order to gain a still clearer understanding of signhood, however, it would be useful to formulate precise definitions. Short's reconstruction of Peirce's theory contains two such definitions, which provide the necessary and sufficient conditions for interpretation and signhood, respectively. Since interpretation is prior to signhood, it is apt to start with the former. Here is a slightly modified version of Short's definition of interpretation, relabeled by me as subjective signification (Short, 2007, p. 158):

#### Subjective signification

*R* interprets *X* as a sign of *O* if and only if

- (a) *R* is, or is a feature of, a response to *X* for a purpose *P*
- (b) *R* is based on a relation, actual or past or apparent or supposed, either
  - i. of *X* to *O*,
  - ii. of *X* to things of *O*'s type,

---

<sup>53</sup> For the record, the two definitions, as formulated by Short, are not encountered in Peirce's own work. They are instead meant to provide a faithful and charitable reconstruction of Peirce's sign theory in the 'mature' stage of its development.

- iii. of things of  $X$ 's type to  $O$ , or
  - iv. of things of  $X$ 's type to things of  $O$ 's type.
- (c)  $O$ 's obtaining has some positive bearing on  $R$ 's appropriateness to  $P$ .<sup>54</sup>

Some clarificatory remarks concerning the three conditions are in order. In accordance with what was said in section 2.1, a purpose, as referred to under (a), is to be understood as a type of outcome which an organism tends to bring about across a variety of circumstances, by virtue of a selected-for disposition to flexibly modify its behavior under certain conditions. Now, as indicated earlier, Peirce held that interpreting responses occur not only in overt behavior, but also in cognition and affect. While perhaps not strictly purposive on the above sense of purposiveness, cognitive or affective interpreting responses qualify as purpose-directed by virtue of being inseparably connected to overt behavior which *is* purpose-directed. Through making possible, and giving rise to goal-directed physical activity, thoughts and feelings constitute auxiliary means to achieving the purposes of overt behavior, and are hence purposeful in an indirect way.

According to the second criterion, (b),  $R$  must be based on a relation between  $X$  and  $O$ , or between instances of either. This can be any kind of relation, ranging from physical proximity ( $X$  being close to  $O$ ), to correlation (instances of  $X$ 's type being frequently followed by instances of  $O$ 's type), to rule-governance (instances of  $X$ 's type being associated with  $O$  via a rule of interpretation), etc. What does it mean for a response  $R$  to be based on some relation? First off, being-based-on is not to be associated with any such thing as causation or counterfactual dependence. Instead, what constitutes  $R$ 's being based on some relation between  $X$  and  $O$ , is simply that relation's role as an actual, past, apparent or supposed practical justifier of  $R$ .<sup>55</sup> In other words, for  $R$  to be based on some relation, that relation must actually, historically, apparently, or suppositionally render  $R$  appropriate as a means to achieving  $P$ . And just as much as  $R$  need not be intentionally produced for it to be purpose-directed,  $R$ 's being based on some relation between  $X$  and  $O$  does not always require that the agent actively considers, or is even remotely aware of, that relation. Only in the case of suppositional justification is there an implication that the interpreting response at issue must involve the exercise of certain cognitive faculties on the part of the agent (e.g. the capacity to entertain the hypothesis of there being a relation between  $X$  and  $O$ ).

---

<sup>54</sup> The original version reads: "R interprets X as a sign of O if and only if (a) R is or is a feature of a response to X for a purpose, P, (b) R is based on a relation, actual or past or apparent or supposed, of X to O or of things of X's type to things of O's type, and (c) O's obtaining has some positive bearing on R's appropriateness to P." My definition thus differs from Short's in the sense that within clause (b), the latter contains only possibilities i and iv. I deemed it necessary to add ii and iii, as I judged the original to miss out on certain clear-cut instances of signification. Apart from this adjustment to the content of Short's definition, I have italicized the variables and arranged the conditions as a list.

<sup>55</sup> I thank dr. Short for pointing this out to me in private correspondence.



## 2 Agents, signs, and symbols

The third condition, (c), dictates that it should minimally be the case that the obtaining of *O* renders *R* at least somewhat appropriate with respect to achieving *P*. In line with what was said in section 2.2, it is not necessary that *P* is fully or even partially achieved in order for *X* to be interpreted by *R* as a sign of *O*. What is necessary is only that there is (actually), was (historically), seems to be (apparently), or would be (suppositionally) at least a marginal increase in the likelihood of achieving *P* through exhibiting *R*, and that *R*, qua means to attaining *P*, is, was, appears to be, or would be sufficiently (cost-)effective in comparison to any available alternatives. The notion of appropriateness is thus relativized to the nature of the justifying relation. Something similar applies to the notion of ‘obtaining’:

‘Obtaining’ is intended to cover a wide variety of conditions, from existing to being a fact to being a possibility; (...) which kind of obtaining obtains depends on how *R* is explained, for that will determine *O*’s ontological category. (Short, 2007, p. 158)

This means that in case the justifying relation is merely past, apparent or supposed, the obtaining of *O* must also be considered as being past, apparent or supposed. The same can be said about nature of the object *O*: “*O* is the object *so far* as it is determined by the relation on which *R* is based, *given R*’s purpose, *P*” (idem, p. 159).<sup>56</sup> Hence, both the nature of *O* and the way in which it obtains, must be considered in light of how the interpreting response *R* is justified.

The above definition of subjective signification determines what it is for something to be interpreted as a sign of something else. This definition, however, only gets us halfway. We would also like to know, after all, what constitutes signification in a more objective sense: under what conditions something *is* a sign of something else. Barring such conditions, we would not be able to account for the intuitively viable distinction between correctly and mistakenly interpreting something as a sign. The sought-after conditions are captured in Short’s second definition (idem, p. 160). With modifications that parallel those made to the first definition, my rendition of that definition is the following:

### Objective signification

*X* is a sign,  $\Sigma$ , of *O*, if and only if there is a relation, either

- i. of *X* to *O*,
- ii. of *X* to things of *O*’s type,
- iii. of things of *X*’s type to *O*, or
- iv. of things of *X*’s type to things of *O*’s type

---

<sup>56</sup> This also implies that there are virtually no ontological limitations on what sort of thing *O* may be. We can countenance anything conceivable or utterable – anything that can function as a target of signification within the cognitive, affective or behavioral bounds of the interpreter. This includes objects that cannot exist, such as the putative object denoted by ‘the square circle’.

such that, for a possible purpose  $P$ ,  $X$  could justifiably be interpreted on the basis of that relation as being a sign of  $O$ .<sup>57</sup>

There are two things to note about the difference between objective and subjective signification. First of all, in case of subjective signification, all that is needed is that there is, was, seems or is taken to be, a relation between  $X$  and  $O$ , or between instances of either, that serves as the basis – the ground – for  $R$ 's interpreting  $X$  as a sign of  $O$ . In other words, the relation may, but need not be, actual. In case of objective signification, however, the basing relation is necessarily actual. Hence, if some  $X$  is a subjective sign of  $O$  in virtue of its being interpreted on the basis of some relation, then this subjective sign only qualifies as an instance of an objective sign ( $\Sigma$ ) of  $O$  if there actually is such a relation. The second thing to note is that the definition of objective signification, unlike the definition of subjective signification, makes no reference to a particular response  $R$ . This is because something is an objective sign not relative to any particular  $R$ , but relative to a set of potential  $R$ 's which would, if manifested, instantiate the interpretation of  $X$  as a sign of  $O$  based on some actual prior relation between the two. These potential responses, although they do not have to be sufficient for attaining  $P$ , must have in common that they contribute at least somewhat to the attainment of  $P$ . The fact that objective signs thus hinge on possible, and not on actual responses, makes it so that objective signification is not a matter of actual interpretation, but of *interpretability* (idem, p. 162). Succinctly put, something is a sign if one would be justified in interpreting it as such relative to a particular purpose.

An example may serve to illustrate how the two definitions work, and how they differ from one another. To use a scenario that is somewhat similar to the deer-example used earlier, consider a situation in which a lumberjack called Sean bolts at the sight of a large tree falling towards him. This scenario can be understood as an instance of subjective signification in the following way. First of all, there is an event, the falling of the tree towards Sean, which figures as the item  $X$ . Let us assume that Sean's purpose ( $P$ ) in this scenario is to avoid severe injury to his body (or worse), and that his act of running away forms the response ( $R$ ) to  $X$  that is directed at this purpose. Although Sean may not have had the time to consciously reflect on it, it is reasonable to say that his response is based on an experienced or assumed correlation between the event-type of a tree falling towards him ( $X$ ) and the occurrence of bodily injury ( $O$ ). So construed, at least,  $O$ 's obtaining has a positive bearing on  $R$ 's appropriateness to  $P$ . That is, Sean's response is subjectively at least somewhat appropriate as a means for avoiding bodily injury, even if in actuality, his

---

<sup>57</sup> The original version, not labelled 'objective signification', reads: "X is a sign, S, of O, if and only if X has such a relation to O, or things of X's type have such a relation to things of O's type, that, for a possible purpose, P, X could justifiably be interpreted on that basis as being a sign of O." I have replaced Short's 'S' with the Greek ' $\Sigma$ ', given that I have already used the letter S for *subject* in section 1.5.

response does not suffice for achieving that goal. The above analysis can be summarized as follows: a tree falling in Sean's direction ( $X$ ) is interpreted through his act of running away ( $R$ ) as a sign of the occurrence of bodily injury ( $O$ ).

Given the possibility of misinterpretation, subjective signs need not always be objective signs. What is required for a subjective sign to be an instance of an objective sign is that the relation upon which  $R$  is based, is indeed an actual one. In case of the above example, this entails that the falling tree is an objective sign of the occurrence of bodily injury only if there actually is a relation between tree-fallings and occurrences of bodily injury.<sup>58</sup> If this is indeed so, Sean's interpreting the falling tree ( $X$ ) as a sign of the occurrence of bodily injury ( $O$ ) through running away ( $R$ ) is an instance of one of the myriad possible ways in which  $X$ 's being a sign ( $\Sigma$ ) of  $O$  can be actualized. Myriad, because there is in principle an unlimited number of ways in which Sean could have manifested  $R$ . Sean might have run at lightning speed or stumblingly, and he might have done so quietly or whilst yelling hysterically – variations that would not have impinged upon the justifiedness of Sean's interpreting  $X$  as a sign of  $O$ . In other words, Sean's response to the falling tree could have taken shape in a variety of manners that would all have been practically justified ways of interpreting  $X$  as a sign of  $O$  relative to his purpose of avoiding bodily injury. Generally speaking, the nature of any objective sign  $\Sigma$  is determined by the  $R$ 's that jointly demarcate  $\Sigma$ 's boundaries of interpretability. Which  $R$ 's are relevant is mostly determined by the characterization of the purpose. On the whole, the more detailed the description of the purpose, the more restricted the set of relevant  $R$ 's will be. As a final note on the implications of the two definitions, it is worth emphasizing that both subjective and objective signification are established independently from the possible availability of comprehending observers. Whether an agent interprets something as a sign is in no way determined by the extent to which a potential onlooker could make sense of her as doing so. In fact, such 'making sense of' is itself an act of interpretation, one that takes another agent's behavior to be a sign of a particular instance of subjective signification.

Let us halt here, and take stock. Drawing on Short's reconstruction of Peirce's sign theory, I have in this section tried to further clarify and specify the account of signification introduced in section 2.2. The definitions presented above provide the starting point for much of what follows. This includes the attempt to define the notion

---

<sup>58</sup> Of course, the fact that certain details were left unspecified makes it difficult to issue any definitive verdict about when that relation can be said to 'actually obtain'. For instance, we do not know what kind of agent we are dealing with, and how 'towards' and 'large' are to be interpreted. It may be that the falling tree is not relevantly correlated with the occurrence of bodily injury, simply because Sean is unusually sturdy, or because the tree is large only in virtue of being comparatively big compared to other specimens of an otherwise tiny kind of tree. Such issues, however, do not need to be ironed out here. I take it that the case description suffices for illustrating the general workings of subjective and objective signification, and for explaining the relation between the two.

of a symbol in section 2.5, but extends also to the next chapters, in which (aspects of) the definitions will at various places be implicitly and explicitly referred to.

#### 2.4 Type- and token-signs

In the previous sections, we saw how Peirce's theory of signs, as reconstructed by Short, could be used to specify the general pattern common to all forms of 'standing for'. This was the first, preparatory step in our inquiry. The next step is to determine which subset within the overarching class of signs is comprised by symbols. In this section, I shall first of all argue that the intended explicatum of symbolic understanding entails that symbols, irrespective of how we wish to characterize them, must be signs that are capable of being instantiated. Hence, symbols must be types. Then, I will formulate definitions of what I call type- and token-signs, and illustrate to what sort of phenomena they apply. Although these definitions, and the explanations that accompany them, amount to a slight departure from my main source material, everything I say in this section (and beyond) is still indebted to Short's reconstruction of Peirce's semeiotic. Passages that are not explicitly referenced, yet which contain claims or terminology reminiscent of Peircean sign theory, will be provided with clarificatory footnotes where necessary. All in all, what follows in the next few pages will be an attempt to cast some of Peirce's thought concerning signification into a system of terms and definitions that lends itself to being exploited for the purpose of defining what symbols are in section 2.5.

As indicated at various occasions already, I have set myself the task to explicate symbolic understanding in terms of the *comparatively more competent use of symbols by an agent*. Now, without introducing substantial new claims about the meaning of this phrase, we can distil from it one tacit assumption. This assumption is that there is something for explicatum-derivatives such as 'in context *C*, symbol  $\Phi$  is used more competently by agent *A* than by agent *B*', to be used *correctly*: there are conditions which determine when it is to be applied, and when not. This assumption should be understood in a way which implicates few, if any, substantial philosophical commitments. To get a feel for this 'thin reading', let me briefly highlight three ways in which the assumption should *not* be understood. First of all, it does not imply that the kind of correctness at stake here should be explicated in terms of the satisfaction of truth conditions. I am here concerned merely with pointing out that the application of an explicatum-derivative is criterion-governed, not with identifying what kind of criteria are involved therein. Hence, the notion of correctness here acts simply as a placeholder concept. Secondly, the claim does not entail that I subscribe to some form of normativism about meaning. Following a view due to Glüer & Wikforss (2015), I maintain that the obtaining of correctness conditions for a linguistic expression amounts merely to there being a particular way of sorting possible uses of that expression into 'correct' and 'incorrect'. From this, normative consequences cannot be derived deductively (*idem*, p. 66). In other words, the claim carries no implication to the effect that there are normative facts about how

an explicatum-derivative *ought* to be used.<sup>59</sup> Thirdly, and finally, I emphasize that the fact that the use of an expression is *criteria-governed* does not entail that any particular instance of use of that expression is also *criteria-guided*: the fact that there is something for explicatum-derivatives to be applied correctly, does not entail that a given user of any such derivative is (consciously) guided in his application by the relevant standard of correctness.

The above assumption thus entails no more than that our explicatum is a general type of expression whose applied derivatives are categorized into ‘correct’ and ‘incorrect’ ones on the basis of certain criteria.<sup>60</sup> While this may sound like a platitude, there is one feature of correctness conditions that is of particular interest. This is that correctness conditions never apply uniquely to only one particular instance, but always (also) to an indefinite number of possible instances. An example may serve to illustrate this. Suppose, for instance, that you watch your neighbor Anna drive a nail into a piece of wood through the use of a hammer. Approving of what you see, you use the expression ‘Anna’s use of hammer *H* in context *C* is competent’ in order to predicate ‘is competent’ of Anna’s hammer-usage. Evidently, there must have been some criterion which governed (though not necessarily guided) your attribution of competence. Let us stipulate that that criterion was the following: ‘an agent’s use of a hammer is competent if and only if an act of hammer-swinging on the part of the agent terminates in a perceptible sinking of the nail into the wood’. It is not hard to think of other scenarios, apart from the one just described, which would have licensed an ascription of competence or incompetence based on that same criterion. For instance, it would not have mattered for the applicability of the criterion if Anna’s hammering had been accompanied by her singing an aria from Wagner’s *Tannhäuser*. In general, there is always a range of possible cases to which a criterion applies. Even if a criterion is highly specific, it will always be possible to conceive of minute alterations of a case description which do not cancel the criterion’s applicability.

This insight has important ramifications for the present inquiry. For if we assume that the application of the explicatum of symbolic understanding is criterion-governed, and if criteria necessarily apply to a range of actual and/or possible cases, then symbols must be signs that are capable of occurring repeatedly.<sup>61</sup> For if they did

---

<sup>59</sup> In principle, this leaves open the possibility of reconciling my view with a ‘pure use theory’ of linguistic meaning (Wikforss, 2001, p. 203), along the lines of the later Wittgenstein. But first and foremost, it allows us to forgo the hefty burden of having to suppose that there are normative facts.

<sup>60</sup> In order to comply with ordinary practice, we should perhaps replace this binary template by a more fine-grained schema that contains categories which are qualified in terms of relative degrees of correctness. But for the argument I develop here, this is not really relevant.

<sup>61</sup> Recall that early on in the second section, I posited that symbols – whatever we take them to be – are signs in the most general sense of that term. Per the definition formulated in the previous section, I here mean by ‘sign’ anything that satisfies the requirements for objective signification.

not have this capacity, the evaluation of their use could not be criterion-governed. But what could ‘occurring repeatedly’ mean, in the case of symbols? That everyday objects can occur repeatedly goes without saying: the chair behind my desk occurred in my experience yesterday, and does so again right now. One might suppose that symbols are likewise repeatable in this most straightforward of senses. However, this would be to lose sight of the nature of signs. Signs, per the definitions formulated in the previous section, only obtain insofar as they form part of a triadic relation with a purpose-directed response and an object. Responses being events, one and the same response cannot occur repeatedly in the sense that one and the same chair can occur repeatedly. Hence, a sign is not something that can be used repeatedly in the way that everyday objects can. The only way in which we can think of a sign as being repeatable, is as something which is capable of being instantiated. In other words, it is only qua *type* that a given sign can be said to occur repeatedly in the form of different *tokens* of that type. Symbols, by implication, must be signs that can be instantiated.

Before we proceed, some terminology needs to be fixed. I use the term *type-sign* to designate items that comprise a subset within the general category of signs. Simply put, type-signs are signs that are types. Furthermore, let us denote another set of items belonging to the category of signs by means of the term *token-sign*. A token-sign is a sign that is a token of a type-sign. Obviously, the categories of type-signs and token-signs are disjoint: no sign that is a token-sign is also a type-sign, and vice versa. Symbols are one kind of type-sign. Symbols such as words and diagrams are type-signs, but so are non-symbolic signs such as medical symptoms. Before we can make an attempt at defining what symbols are, we need to define the overarching classes of type-signs and token-signs. Admittedly, carrying out this task in full would require entering into a complex ontological debate surrounding the type-token distinction.<sup>62</sup> Since we need to make headway in what is still an epistemological inquiry, I will not attempt to provide a full-fledged account of what types and tokens are.<sup>63</sup> Instead, I will make do with the following platitude:

*Types and tokens*

A token is an item with a unique spatio-temporal location that stands in a relation of instantiation to a type, which lacks a definite spatio-temporal location but which is capable of being instantiated by tokens.<sup>64</sup>

---

<sup>62</sup> Given that the type-token distinction is in fact a Peircean invention, this might also require diving knee-deep into Peirce’s phanerescopy. As noted in section 2.2, I want to steer clear of this subject matter, despite its mysterious lure.

<sup>63</sup> Nor will I here attempt to allay nominalistic qualms regarding type-token talk in general.

<sup>64</sup> There is one further distinction worthy of mentioning here. According to Wetzel (2018), tokens of types should be distinguished from *occurrences* of types. The difference can be elucidated as follows. Somewhat akin to the way in which a hammer may occur repeatedly in various acts or events, a type may occur repeatedly in another type in which it participates. For instance, the type *digit* occurs five times in the type *human hand*. Importantly, these five

## 2 Agents, signs, and symbols

According to this definition, a type-sign must be a kind of sign that lacks a definite spatio-temporal location, but which is capable of being instantiated by token-signs.

Let us now specify in a bit more detail what type- and token-signs are, starting with the latter. We established earlier that something is a sign if and only if it is interpretable as such relative to some purpose of interpretation. We therefore first need to know what it is for something to be *interpreted as* a token-sign in order to determine what it is for something to *be* a token-sign. Recall from the previous section that the notion of a subjective sign (i.e. being interpreted as a sign) comes with the following definition:

### Subjective signification

*R* interprets *X* as a sign of *O* if and only if

- (a) *R* is, or is a feature of, a response to *X* for a purpose *P*
- (b) *R* is based on a relation, actual or past or apparent or supposed, either
  - i. of *X* to *O*,
  - ii. of *X* to things of *O*'s type,
  - iii. of things of *X*'s type to *O*, or
  - iv. of things of *X*'s type to things of *O*'s type
- (c) *O*'s obtaining has some positive bearing on *R*'s appropriateness to *P*.

What is required for something to be interpreted as a token-sign? Given that tokens are defined as instances of types, this defining aspect of tokens must figure somehow in the prior relation upon which a given *X*'s interpretation as a token-sign is based. I thus propose the following: for some *X* to be interpreted as a token-sign of some *O*, the response *R* to *X* must be based on a relation that obtains (actually, historically, apparently or suppositionally) between instances of *X*'s type and *O*, or between instances of *X*'s type and instances of *O*'s type. In order to adjust the above definition so that it applies only to token-signs, we therefore need only delete i and ii from clause (b).<sup>65</sup> The definition of subjective token-signification therefore looks as follows:

### Subjective token-signification

*R* interprets *X* as a token-sign of *O* if and only if

- (a) *R* is, or is a feature of, a response to *X* for a purpose *P*
- (b) *R* is based on a relation, actual or past or apparent or supposed, either
  - i. of things of *X*'s type to *O*, or
  - ii. of things of *X*'s type to things of *O*'s type

---

occurrences are not separate items with distinct spatio-temporal locations. Rather, the five digit-occurrences in the human hand are of the same item: the digit in general.

<sup>65</sup> In footnote 69, I use an example to illustrate the difference between signs that are based on a relation between particulars simpliciter, and signs that are based on a relation between particulars qua instances of types.

- (c)  $O$ 's obtaining has some positive bearing on  $R$ 's appropriateness to  $P$ .

Given what was said about objective signification in the previous section, the above yields the following definition of objective token-signification, where token-signs are denoted by ' $\Sigma_i$ ':

Objective token-signification

$X$  is a token-sign,  $\Sigma_i$ , of  $O$ , if and only if there is a relation, either

- i. of things of  $X$ 's type to  $O$ , or
- ii. of things of  $X$ 's type to things of  $O$ 's type

such that, for a possible purpose  $P$ ,  $X$  could justifiably be interpreted on the basis of that relation as being a sign of  $O$ .

Having established what token-signs are, it would only seem a short route towards providing a definition of type-signs. However, there is a complication. According to Peirce's theory, signs are response-dependent: something only ever is a sign relative to actual and/or possible responses which render it so. Clearly, responses must come from (possible or actual) agents situated in space-time, which means that whatever is responded to had better be situated in space-time as well. Thus emerges the following problem: if type-signs lack definite spatio-temporal locations, it is difficult to see how they could induce responses in agents. And if something cannot induce a response, it cannot operate as a sign. Hence, in contradiction with my earlier characterization of type-signs as being signs that are types, it appears that type-signs (including symbols) are not really signs at all. At best, they would be signs only in the derivative sense of having tokens that signify. Should we concede this point, and accept that type-signs are not genuine signs in the sense of Short's definitions? This is a difficult issue, the resolution of which would ultimately require entering into the ontological debate over the nature of types and tokens. In what follows, I shall adopt a middle course that avoids explicit commitment to the idea that types can be signs, but which does recognize the notion of the type-sign as a legitimate object of thought and discourse.<sup>66</sup> The following definition of type-signs is meant to specify this knife-edge position:

---

<sup>66</sup> This is one respect in which I deviate from Peirce. Peirce held that generality, or 'lawfulness', could truly enter into our experience, and consequently that types could genuinely signify. As Short (2007, p. 81) explains: "[L]ike Kant and unlike the British empiricists, Peirce held that thinking enters into basic forms of experience. (...) [U]nlike both Kant and the empiricists, Peirce held that the content of sense experience, prior to analysis, is itself continuous. (...) The continuous whole has to be given first, before it can be analyzed as divisible." For Peirce, generality can be responded to in virtue of the fact that thought impinges directly upon our experience of a spatio-temporally continuous whole, or 'phaneron'. This is an intriguing view, but also a controversial one. Given the fact that I here resist defining the notion of a type-sign in terms of objective signification, I am effectively



*Type-sign*

A type  $T$  is a type-sign,  $T_{\Sigma}$ , if and only if for all tokens  $t$  of  $T$ , some object  $O$ , and some relation  $L$ ,  $t$  is a token-sign ( $\Sigma_t$ ) of  $O$  that has  $L$  as its basis.

This definition does not characterize the type-sign in terms of the definition of objective signification, thereby making room for the position that type-signs are not really signs. Instead of delineating a particular category within the general class of signs, it delineates a particular category within the general class of types. At the same time, and perhaps misleadingly, it does retain the term ‘type-sign’. Let me briefly explain why I think continued usage of that term is permissible. For some item to be a token-sign, it must be the case that the possible interpretation of that item as something which signifies *qua* instance of some type-sign, is justified relative to the purpose of that interpretation. Only through being so interpretable can an item constitute a token-sign. There is thus a mutual dependency between token-signs and their associated type-signs: types do not signify except through their tokens, but tokens do not signify except through being interpretable as instances of type-signs. If we were to disavow talk of type-signs altogether, this might incorrectly suggest that the signifiatory role of token-signs is somehow prior to, or independent from, the types they instantiate. For this reason, I will continue to speak freely of signs and symbols in both their type- and token-senses, entrusting to the reader the task of adding the qualifications just discussed.

Now for some examples. Where do we encounter type-signs and their tokens? A quick survey confirms that everyday experience is rife with them. For instance, the inscription ‘**WARNING**’ on a road sign signifies something *qua* word insofar as it is interpretable as an instance of a general type, namely the type-sign which has as its tokens the innumerable possible inscriptions of (the imperative use of) the word *warning*. The inscription ‘**WARNING**’ operates as a token-sign of that type-sign in virtue of its possible interpretation being grounded in a relation that instances of the inscription’s type bear to an object, or to instances of a type of object.<sup>67</sup> Generally speaking, all words of natural language are type-signs that are capable of being instantiated by token-signs in speech, writing, and/or manual signs. More generally still, any particular gesture, mark, or sound which signifies in virtue of some explicit or implicit convention, does so through instantiating a type-sign. But the occurrence of type- and token-signs is by no means limited to human affairs. The event-type of swiftly gathering clouds is associated with the possibility of a future downpour, and as such functions as a type-sign instantiable by clouds gathering swiftly at a

---

siding with a more conservative position, according to which generality is something that is posterior to our immediate experience.

<sup>67</sup> The object, in this case, is a request of sorts, which, in conjunction with some other sign (e.g. pictograms), asks the road user to exercise caution with respect to some actual or potential danger (e.g. road works, or the crossing of wild animals).

particular place and time.<sup>68</sup> Again, generally speaking, the interpretation of evidence and the making of predictions is a matter of items (subjectively and possibly objectively) signifying as tokens of types. Type- and token-signs are thus ubiquitous in both communication and in explanation and prediction.<sup>69</sup> In what remains of this chapter, I will delve deeper into the nature of those type- and token-signs that constitute our means for communication: symbols and replicative signs.

## 2.5 Symbols and replicative signs

Having seen what type- and token-signs are, let us now turn to answering this chapter's main question concerning the nature of symbols. So far, we have established that symbols must be type-signs.<sup>70</sup> What subset within the category of type-signs should the term 'symbol' denote? This depends on what sort of things our explicatum is supposed to apply to. Recall from the previous chapter that symbolic understanding is associated not just with the understanding of linguistic expressions,

---

<sup>68</sup> Think also of our earlier example about Sean making his way to safety in response to the sight of a falling tree: the falling tree signified the occurrence of bodily harm to Sean, because of that sign being grounded in a correlation between instances of the type 'tree falling in close proximity to a person' and the type 'occurrence of bodily harm'.

<sup>69</sup> Given the ubiquity of type- and token-signification, one might wonder whether there is any room left for signs to signify 'by themselves', rather than qua instance of some type. Let us call this putative kind of sign a *singular sign*. To understand why I think singular signs do indeed exist, we have to consider what the definition of objective singular signification would look like. First of all, the definition for objective token-signification dictates that in order for an  $X$  to be a token-sign of some  $O$ , it must be interpretable as a sign on the basis of a relation between instances of  $X$ 's type and  $O$  or between instances of  $X$ 's type and instances of  $O$ 's type. In case of objective singular signification, on the other hand, the first relatum of the justifying relation would have to be  $X$ , simpliciter. Now, we can think of instances of objective signs which are justified based on a relation that has  $X$  as the first relatum. Think, for example, of instinctively retracting your hand at the touch of a hot pan. The touch of the hot pan is  $X$ , the instinctive retraction of your hand is  $R$  and  $O$ , presumably, is the onset of a burn.  $R$ , in this case, need not be based on a relation obtaining generally between hot-pan-touchings and burn-onsets. Instead, it may be based on a relation which obtains between  $X$  and  $O$  simpliciter. The touch of a particular item – this hot pan right here – is followed by a particular event – the onset of a burn, right now – and it is the relation of immediate temporal succession that obtains between these two particulars which can serve as a practical justifier of  $R$  relative to the purpose of avoiding burns. To be sure, one may, whilst retracting one's hand, entertain a thought which does interpret the touching of the pan as a token-sign. But insofar as  $R$  occurs without premeditation, such a thought would be an additional response to  $X$ , independent from  $R$ . [Addendum: those familiar with Peirce's sign theory will probably be inclined to associate the notion of a singular sign with Peirce's concept of a *sinsign*. Indeed, the fact that the prefix 'sin' means 'being only once' (*EP* 2:291) would seem to confirm this idea. However, I think that the two concepts are not wholly commensurable, let alone coextensive. For Peirce, tokens of types are always sinsigns. Consequently, insofar as types signify through their tokens, they must do so via sinsigns. By contrast, my proposal implies that the categories of token-signs and singular signs are disjoint: no token-sign is a singular sign, or vice versa.]

<sup>70</sup> Of course, we should continue to bear in mind the proviso regarding type-signs: type-signs are not really signs, but rather types whose tokens are signs.

but also with the understanding of non-linguistic representational devices. Thus, the understanding of such a wide array of items as words, sentences, maps, diagrams and physical models, falls within the perimeter of symbolic understanding. But if all these items are symbols, or systems of symbols, then what could be their common denominator? Drawing once more on Short's interpretation of Peirce (Short, 2007, p. 210-14), I suggest this is the following: what connects the various kinds just mentioned, is not just the fact that their tokens signify, but rather the fact that their tokens signify through being generated or adduced *in order to signify*. A symbol is thus a special sort of type-sign, namely one whose tokens have a significatory purpose.<sup>71</sup>

How to account for this kind of sign in terms of Short's definitions? For starters, we need to allot tokens of symbols their own label. The word 'replica', another item from Peirce's terminological repertoire, fits the bill nicely: replication, as a form of instantiation, implies a purpose to instantiate. However, because replicas as such need not be signs (think of today's 234<sup>th</sup> Mini Cooper S rolling off the production line), let us call a token of a symbol a *replicative sign*. Respecting the priority of the subjective sign with respect to the objective sign, we first need to determine what it takes for something to be interpreted as a replicative sign. Crucially, a definition of subjective replicative signification – apologies for the tongue twister – needs to capture the idea that for *X* to be interpreted as a replicative sign, it must be interpreted as a token-sign that is 'there' to signify. This means that for some item *X* to be a subjective replicative sign, there must be a response *R* to *X* that is based on a particular kind of relation between things of *X*'s type and *O* or things of *O*'s type – a relation which is such that it invests tokens of *X*'s type with the purpose to signify *O* or things of *O*'s type. I call such a relation a *rule of interpretation*, and I stipulate such rules to be conditional imperatives:

*Rule of interpretation*

A rule of interpretation is a conditional imperative of the form 'If a token of *X*'s type is encountered, interpret it as a sign of *O*, or of things of *O*'s type'.

---

<sup>71</sup> This harks back to Augustine's distinction between *signa naturalia* and *signa data*, which was briefly mentioned in section 2. The category of *signa data* (that is, signs that are there to signify) occurs also in Peirce's work, although Peirce did not demarcate the category consistently. As Short (2007, 210) points out, Peirce's notion of a so-called *legisign* wavered between one that included the class of *signa data* as a subset, and one that was coextensive with it. A *legisign*, according to its most general definition, is "a law that is a sign" (*EP* 2:291). A law, in Peirce's sense, can be a generality of any sort, including laws of nature. This seems to suggest that *legisigns* comprise all sorts of significatory generalities: *signa data* as well as *signa naturalia*. In the same breath, however, Peirce adds that such a law is "usually established by men" and that it is "a general type which, it has been agreed, shall be significant" (*ibidem*). This is more in line with an interpretation according to which 'legisign' is synonymous with 'signum datum'. There is no need to settle this issue of interpretation here, but it does perhaps aid the understanding of the Peirce-informed reader to know that what I call a symbol more or less corresponds to the restrictive conception of the *legisign*.

Why should signficatory purposes be grounded in interpretative rules, though? The reason is that to invest an item with a purpose is to delimit how the item is to be conceived of, or how it ought to be used. By way of illustration, consider the example of a landowner who drives a wooden pole into the ground (or more plausibly: instructs someone else to do this for him), saying: “Let this pole mark the outer boundary of my territory”. Without this performative speech act, the pole would have just stood there, serenely void of purpose. The landowner’s declaration assigns a purpose to the item, *via* the imposition of an interpretative rule. Thus, generally speaking, to invest an item *X* with a signficatory purpose is to subsume it under a rule which says: if you encounter *X*, interpret it (via a response *R*) as a sign of some *O*, or of things of *O*’s type.

Some more needs to be said about subjective replicative signification before I can formulate its definition. Recall that in order for an item *X* to be interpreted, through a response *R*, as a sign of an object *O*, there must be an actual, historical, apparent or supposed prior relation between *X* and *O* which justifies *R* relative to its purpose *P*. Now, if *R* is to be based on a rule in the sense of that rule justifying (in one of the four modes) *R* relative to *P*, then it must be the case that the rule renders *R* appropriate as a means for achieving *P*. But relative to what sort of purpose does a rule constitute a justificatory basis for an interpreting response?

The answer, I take it, is: relative to the purpose of attaining the outcomes associated with compliance with the rule. For example: if I deem a particular wooden pole to be governed by a rule according to which the pole is a boundary marker, then this rule subjectively justifies my response to (an encounter with) the pole in case that response is aimed at realizing or anticipating outcomes of the rule’s being operative. Let us suppose that one such outcome, in this case, is the incurrance of a financial penalty upon transgression of the boundary. If I happen to be walking in the direction of the pole, but, upon sighting it, resolve not to venture beyond the pole in order to avoid incurring the penalty associated with the pole’s purported function, then the interpretative rule upon which this response (i.e. the decision not to venture beyond the pole) is based justifies that response relative to its purpose of avoiding a certain consequence of non-compliance with the rule.

Let me now try to formulate a definition of subjective replicative signification. I just argued that replicative signs form a distinctive category among the larger class of token-signs because they have a purpose to signify, and that the notion of a signficatory purpose implicates the idea of a rule of interpretation. Based on this, I further showed that if an item *X* is to be interpreted as a replicative sign via a response *R*, then *R* must be directed at the purpose of complying with the rule which is taken to govern *X*. For only through being directed at complying with a rule can a response be said to be practically justified by that rule.<sup>72</sup> In summary, then, a subjective

---

<sup>72</sup> Given that we are still only dealing with *subjective* replicative signification, one should keep in mind that the rule is, in a sense, self-enforced. The agent, in responding to *X*,

replicative sign is a token-sign that is based on a rule of interpretation, and which operates relative to a purpose of complying with that rule. In more orderly fashion:

Subjective replicative signification

$R$  interprets  $X$  as a replicative sign of  $O$  if and only if

- (a)  $R$  is, or is a feature of, a response to  $X$  for the purpose of complying with the rule of interpretation cited under (b)
- (b)  $R$  is based on a rule of interpretation, actual or past or apparent or supposed, which says that things of  $X$ 's type are to be interpreted as signifying  $O$ , or as signifying things of  $O$ 's type
- (c)  $O$ 's obtaining has some positive bearing on  $R$ 's appropriateness to  $P$ .<sup>73</sup>

Obviously, not all ways of interpreting items as replicative signs will amount to correct interpretations. I may interpret my neighbor's utterance 'Help!' as a distress call and come to his aid swiftly, only to discover that he was enthusiastically referring to a song by the Beatles. In short, not just any rule of interpretation that an interpreter takes to be operative, needs to be actually operative. Some  $X$  is an objective replicative sign of  $O$  only in case there are possible acts of interpreting  $X$  (that is, possible  $R$ 's) which are based upon a rule of interpretation that actually applies. Hence:

Objective replicative signification

$X$  is a replicative sign,  $\phi$ , of  $O$ , if and only if

Things of  $X$ 's type are related to  $O$  or to things of  $O$ 's type through a rule of interpretation, such that, for the possible purpose of complying with that rule,  $X$  could justifiably be interpreted on that basis as being a sign of  $O$ .

Replicative signs, being token-signs, instantiate type-signs. These type-signs, we have agreed, are called symbols. As we have learnt from the previous section, type-signs cannot really be signs, since they cannot be responded to. So the notion of a

---

interprets that  $X$  as a replicative sign by virtue of his response being directed at complying with a rule. This should not be taken to mean, however, that any subjective replicative sign is the product of conscious compliance. Recognition of, and compliance with rules, can be a matter of sheer habituation manifesting itself outside of our awareness.

<sup>73</sup> In this definition, condition (c) appears to be redundant. For if  $R$  is directed at the purpose of complying with the rule upon which it is based, then it cannot but be the case that  $O$ 's obtaining has some positive bearing on  $R$ 's appropriateness to that purpose. This, I think, aligns with our intuitions: if I interpret this sentence with the aim of interpreting it in accordance with some rule I take to be operative, then it cannot be otherwise than that the obtaining of the rule's prescription makes my act of interpretation appropriate to its purpose. I have refrained from removing (c), however, in order to maintain a neat parallelism with the previous definitions.

symbol has to be characterized not in terms of some version of the definition of objective signification, but rather as follows:

*Symbol*

A type-sign  $T_\Sigma$  is a symbol,  $\Phi$ , if and only if for all token-signs  $\Sigma_t$  belonging to  $T_\Sigma$ , some object  $O$ , and some rule of interpretation  $I$ ,  $\Sigma_t$  is a replicative sign ( $\varphi$ ) of  $O$  that has  $I$  as its basis.

Finally, then, we have determined what a symbol is: it is a type-sign whose tokens are interpretable as signs on the basis of them being related to an object via a rule of interpretation. Note that I use the  $\Phi$  and  $\varphi$  to refer to symbols and replicative signs, respectively, where considerations of notational consistency would have obliged me to use, for instance,  $T_\Phi$  and  $\Sigma_\varphi$ . I have taken the liberty to ignore these considerations, and instead prioritize concerns of practicality and legibility.

Some residual questions stand in need of answering. First of all, what determines whether some rule of interpretation actually applies? This depends on whether  $X$  falls within the set of items that is related to  $O$  via the rule of interpretation. And this, in turn, depends on whether  $X$  satisfies the specific entry requirements for membership of that set. Those entry requirements do not solely have to do with  $X$ 's observable characteristics. For instance, replicas of the same symbol typically resemble each other closely, yet similarity is no guarantee of replication. The word 'fire' signifies fire through a rule of interpretation, but a smoke column accidentally taking the shape of the word 'fire' does not.<sup>74</sup> Furthermore, two items may be replicas of the same symbol even if each of them is more similar to replicas of other symbols than they are to each other. Two handwritten marks produced by different individuals may closely resemble a question mark and a nine respectively, yet they can operate as replicas of the same sign: the F-clef of western musical notation. As Short (2007, p. 212) points out, replication is effected not by similarity alone, but rather by the exploitation of similarity for the purpose of instantiation. In other words, for a rule of interpretation to be in force between  $X$  and  $O$ , the item  $X$  must be generated, selected or otherwise 'be there' in order to let it instantiate the relevant type via its similarity to other replicas. Hence, similarity does play a role in replication, but only insofar as it is implicated in a purpose to instantiate.

Replicative signs are items which are interpretable as token-signs in virtue of the fact that they are related to an object  $O$  by means of a rule of interpretation. And they are so related just in case they are generated or selected in order to instantiate the set of items to which the rule applies. This gives rise to a further question. For when, exactly, is it the case that something is generated or selected in order to signify? This

---

<sup>74</sup> Of course, the smoke column may signify fire through constituting evidence of it. But in that case it signifies not as a replicative sign that is based on some rule of interpretation, but rather as a normal token-sign that is based on, for instance, fire's propensity for producing smoke columns.

question can be answered in terms of the account of purposiveness laid out in section 2.1. There it was said that a purpose is a type of outcome which an organism tends to realize consistently, through having a disposition to flexibly modify its behavior across differing circumstances. When an agent generates or selects an item in order to let it signify something, what the agent purports to realize are certain effects associated with the sign's rule of interpretation. For instance, the waving of my hand signifies my readiness to part amicably with someone (and my intention to convey this attitude to the recipient), in case my hand-waving is part of a modal pattern of behaviors whose common type of outcome consists in a certain 'standard scenario' associated with parting amicably. Such a scenario would include the exchange of good-byes and the reciprocation of gestures and facial expressions, accompanied with mutually accepted physical separation. In case the act of hand-waving is not part of a pattern of behaviors which lead to this kind of standard scenario, the hand wave is not a parting gesture. When, for instance, I wave my hand in order to shoo a fly, the pattern of behaviors to which that act belongs is associated with a wholly different type of outcome: the fly being successfully shooed.<sup>75</sup>

Since the above account of significatory purposes is grounded upon a non-mentalistic picture of purposiveness, it can accommodate examples of replicative signification in organisms which are not obviously capable of forming, and acting on, intentions. Short (2007, p. 211) mentions the example of mating displays in birds: male individuals of certain avian species, whose behavior we may suppose is instinctual rather than intentional, are wont to produce intricate sequences of movements in order to signify their willingness and fitness to mate to female individuals of that species. In light of what was just said, such mating displays can be regarded as replicative signs, given that they are produced for the purpose of being interpreted in a certain way. For even though the bird itself might not engage in mating behavior with an intention to elicit a certain kind of response, the mating display nonetheless occurs for such a purpose by virtue of the fact that in case of failure, the male bird is disposed to modify its behavior in such a way that the likelihood of success (i.e. his willingness and fitness to mate being successfully conveyed) is increased. For example, the male might reposition himself if it turns out his dance is not properly visible from where the female is located. Or, in case the female flies off in apparent indifference, the male may try putting up a more impressive performance. An important implication of my proposal is thus that

---

<sup>75</sup> What about the significatory purposes of artefacts, such as those of a desktop computer? I contend that those are merely apparent. When your computer instructs you to 'press any key', for instance, it is not the device itself which has the purpose of letting that phrase signify a particular directive. Rather, it is the purpose of the device's manufacturer which ultimately determines by means of which rule of interpretation the expression is meant to be interpreted. In general, significatory purposes of artefacts are strictly speaking only 'quasi-purposes' which ultimately derive from *actual* purposes the manufacturer(s) had in imbuing their creations with certain dispositions.

symbols and their replicas transcend the alleged divide between humans and non-humans.

### **2.6 Conclusion**

Let me now very briefly summarize this chapter's findings, and look forward to what comes next. First of all, I spelled out the notion of signification (i.e. 'standing for') in terms of two definitions borrowed from Short's reconstruction of Peirce's semeiotic. Then, definitions of token- and type-signs were formulated, based on the idea that criterion-governed sign-usage can only pertain to signs that can occur repeatedly. Finally, we found that tokens of symbols, called replicative signs, are distinguishable from other token-signs by virtue of the fact that they have signficatory purposes. This is an important first step towards achieving one of the main aims of this project, which is to provide an explication of symbolic understanding in terms of comparatively superior symbol-usage. For instead of having only an indeterminate list of what we usually take to be symbolic devices (words, diagrams, models, etc.), we now have a definition at our disposal. This definition captures the items in the list, yet provides clearer boundaries to the notion of a symbol than any non-exhaustive list of examples could. In keeping with the principles of Carnapian explication, it thus achieves some measure of exactness whilst retaining considerable similarity to ordinary usage. In the next chapter, I will flesh out the remainder of my explicatum of symbolic understanding. That is, I shall explain how the notion of *competent symbol-usage* is to be understood.





## Chapter 3: Symbolic understanding

Throughout chapter 1, we had to make do with a very rough idea of what sort of things symbolic understanding can take as its object. According to that rough idea, symbolic understanding pertains to a broad and vaguely delineated set of items which includes, among many other things, words, sentences, diagrams, gestures, and models. Chapter 2 has enabled us to replace this loose conception of the domain of symbolic understanding with a definition of symbolicity. That definition tells us that symbols are types, whose tokens are signs that are based on a rule of interpretation. In this chapter, I will attempt to spell out what is required for symbols, so defined, to be used competently. The claim I intend to argue for is that one's symbol-usage can count as competent only if one possesses a disposition for proper symbol-usage, the presence of which is due to one's having engaged in a learning process. Obviously, various notions will have to be clarified over the course of this argument. Among other things, I will attempt to make clear what it means to use a symbol, how a symbol is used properly, and in what sense learning processes bridge the gap between propriety and competence in symbol-usage. In the chapter's final sections, I gradualize the definition of competent symbol-usage, thereby completing my account of what it means for someone to have a better understanding of a symbol than someone else.

The chapter will be structured as follows. In **section 1**, I start out by explaining that the methodological principle I have adopted in section 1.5 prompts us to conceive of symbolic understanding in terms of behavior, and more specifically, in terms of symbol-usage. Then, in an effort to connect the notion of symbol-usage to the account of agency presented in section 2.1, I propose that using a symbol consists in its being selected as a means for a purpose. In **section 2**, I define proper symbol-usage as a specific type of symbol-usage: namely the kind of symbol-usage that consists in a symbol being selected for the purpose associated with the symbol. **Section 3** subsequently explores the question whether propriety is sufficient for competence. I argue that it is not, given that a disposition for proper symbol-usage may be acquired in ways that are tangential to the general aim of ascriptions of symbolic understanding. Based on this finding, **section 4** then suggests that competence requires learning: an agent qualifies for an ascription of symbol-using competence in case she has a disposition for proper symbol-usage which was acquired through the operation of a learning mechanism. I further argue that such a learning mechanism, in turn, can only become operative provided the agent attends selectively to symbolically relevant features of the replicative signs she encounters. At the end of the section, a definition of competent symbol-usage is formulated. In a slight digression from the main topic, I subsequently devote **section 5** to the topic of selective attention, offering an account of what it is and a conjectural explanation of its workings. In **sections 6** and **7** I finalize my explication of symbolic

### 3 Symbolic understanding

understanding, by spelling out what it means to understand a symbol *better* than someone else. I identify two dimensions of graduality, one relating to the comprehensiveness of an agent's disposition for competent symbol-usage (breadth); the other pertaining to the way in which such a disposition is formed (depth).

#### 3.1 How to use symbols

The explicatum of symbolic understanding is meant to retain an important aspect of our pretheoretic notion of that concept, whilst also significantly deviating from it. On the one hand, it preserves our intuition that symbolic understanding is a matter of dispositions, rather than of occurrences. Yet on the other hand, the explicatum differs from its associated explicandum through disregarding the purported cognitive dimension of symbolic understanding. That is, the concept of competent symbol-usage applies only to behavior: using symbols is something which happens in speaking, writing, and gesturing, or more generally in physically influencing and manipulating our environment. Taken together, the commitments to dispositionalism and behaviorism entail that symbol-using competence is ascribed to an agent on the basis of the symbol-using behavior which that agent is capable of manifesting. In this chapter, I will seek to determine under what conditions such ascriptions are warranted. As part of carrying out this task, this starting section will address the question what it means to use a symbol. To that end, I shall explain what it is to use something (whether a symbol or something else). I submit that usage consists in *selection*. Then, I shall explain that symbol-usage involves selection of a peculiar sort: namely one where selection is from a range of possible, rather than actual alternatives. Before I do all this, however, I need to justify my choice for a behaviorist account of symbolic understanding. For why should we want to think of symbolic understanding in terms of *symbol-usage* in the first place?

In section 1.5, understanding-ascriptions were assumed to have the general function of promoting the possession of certain dispositions. I added to this general assumption the following more specific, but as of yet unsubstantiated claim regarding the function of ascriptions of symbolic understanding:

*The function of ascriptions of symbolic understanding*

An ascription of symbolic understanding to an individual serves to encourage that individual to sustain the possession of a disposition that contributes to maintaining a stable practice of symbol-usage within a community.

Now is the time to provide arguments for this claim. Below, I show that a community (any community) needs a normative instrument for maintaining stability in the community's practice of symbol-usage, because communal living requires efficient communication, and the possibility of efficient communication is predicated on the existence of a stable practice of symbol-usage. The concept of symbolic understanding can be seen to cater to the need for preserving a community's practice

of symbol-usage. As I also explain, the concept of symbolic understanding does *not* cater to any purported need for micro-managing the cognitive dispositions (e.g. thoughts, beliefs) of individuals, simply because there is no such need. Only in a very general and indirect sense do ascriptions of symbolic understanding pertain to cognition: namely through promoting certain mechanisms for acquiring dispositions for symbol-usage.

What is efficient communication, and why should a community want it? To see just how efficient our own communicative habits actually are, consider the following everyday example. A host, Mary, welcomes an invited guest, John, to a party by enthusiastically saying ‘hey John!’. Without effort, and probably without even so much as conscious awareness, John will immediately interpret Mary’s greeting as having been intended to convey acknowledgement of John’s status as an invitee, and as an expression of Mary’s willingness – eagerness, perhaps – to have John join the party. John, in turn, will then likely communicate this interpretation of Mary’s greeting by responding to her in similarly enthusiastic fashion, for instance by saying ‘hi Mary!’. If executed properly, John’s utterance at once successfully informs Mary of John’s recognition of Mary’s intentions in her saying ‘hey John!’, and of John’s gladness about his imminent joining of the party. Mutually reassured, the two may exchange a final pair of smiles, John will walk into the building, remove his coat and smoothly merge into the partying crowd. And all of this could happen within a matter of seconds.

Crucially, however, *none* of this could come about without there being more or less stable interpretative rules underlying the use of the words ‘hello’ and ‘hi’, the vocative use of proper names, and the use of raised eyebrows, warm smiles and excited tones of voice. Absent such stability, agents would have a hard time gauging what their fellow community members were up to, as they would have to fall back on ‘hard thinking’ in trying to decipher the interpretative rules their peers were relying on. In such a scenario, communication would be overly time- and energy-consuming, and prone to breed misunderstandings. Indeed, the very structural integrity of a community would be at stake if its means of communication were volatile in this way. For in such a scenario, individuals would be constantly occupied with securing the very foundations of their community, rather than with communal living as such. At some point, they would simply *have* to reach implicit mutual agreements as to how communicative acts should be interpreted in order to have their hands free to actually operate as a community; to wit, by collectively pursuing common interests, by managing mutual dependency relationships, by engaging in productive competition, and so on.

Thus, to the extent that communal living requires efficient communication, interpretative rules must remain more or less fixed. To achieve this, a community must have a means to promote fixity. And this, I take it, is what ascriptions of symbolic understanding are for: they serve to encourage the possession of dispositions for symbol-usage that accords with existent rules of interpretation. Such

### 3 Symbolic understanding

ascriptions pertain to behavior, because it is ultimately the doings of agents that a community needs to coordinate, not necessarily also their inner lives. In and of themselves, cognitive dispositions of individuals do not directly impact on a community's proper functioning. Admittedly, it is hard to imagine that agents can, for example, collectively pursue a shared goal without there being cognitive concordance among them. But a community does not need to exert direct and specific control over its members' internal states in order to function properly as a community.

In fact, the use of such a micro-level tool could turn out to be counterproductive with regard to the goal of maintaining communicative efficiency. After all, different agents may rely on a wide array of cognitive dispositions in acquiring dispositions for symbol-usage. Were a community to preferentially reward only one, or a few, of such possible dispositions, this might undermine the communicative abilities of those who tend to more effectively acquire dispositions for symbol-usage through other cognitive avenues. As I argue in sections 3.3 and 3.4, communities merely need a very general instrument to promote certain cognitive *mechanisms* for generating behavioral dispositions: namely a criterion which says that dispositions for symbol-usage should be grounded in learning. The addition of this criterion does not negate the behaviorist character of my account of symbolic understanding, but instead imposes a further restriction upon what kinds of behavioral dispositions can license ascriptions of symbolic understanding.

That we nonetheless have the intuition that symbolic understanding pertains directly to specific cognitive dispositions is because those dispositions will tend to be affected by whatever external pressure is exerted upon our behavioral habits. If someone is denied understanding of some symbol by his peers, that denial will, from the peers' perspective, pertain to some behavioral mistake he made, and it will have the aim of motivating him to adjust his behavioral disposition(s). Because in the end, the community just needs him to be disposed to *do* the right thing – what he is disposed to believe or feel is not really in the interest of the community as such. But to the failing understander, it will seem as if his peers are in one stroke trying to control both his behavioral *and* his cognitive dispositions, since *he* is reflectively aware that his cognitive dispositions are tied to his behavioral dispositions. As I see it, this is why the pretheoretic notion of symbolic understanding, based as it is upon a kind of 'inward reflection' upon what ascriptions of symbolic understanding do to us, has a behavioral as well as a cognitive dimension. When, however, we adopt a more outwardly directed perspective at what goal ascriptions of symbolic understanding actually serve to a community, the cognitive component all but disappears. Hence my contention that symbolic understanding consists in the *use* of symbols.

This brings me to the next question of this section. What is it to use something, generally speaking? The answering of such a question might seem to be of interest only to the overzealous analytic philosopher. Indeed, as far as I am aware, no

attempts have been made at spelling out the notion of use in terms of necessary and sufficient conditions. And admittedly, there seems to be little demand for any such undertaking. The lack of both precedents and incentives would seem to provide sufficient reason to refrain from dissecting everyday language. Hence, in what follows, I will not try to offer a conceptual analysis of usage. Instead, I shall once more utilize the method of Carnapian explication. The purpose of explicating the concept of usage will be strictly theory-internal: I shall argue that the proposed explicatum, according to which usage is selection, enables us to connect the notion of symbol-usage with the account of agency laid out in section 2.1. In this way, the internal coherence of my theory is enhanced. In order to avoid being reproached for arbitrarily changing the subject, I argue that despite appearances to the contrary, the explicatum captures our everyday sense of the word ‘use’ rather well. The idea is that this legitimizes the choice for a coherence-enhancing departure from the default option of leaving ordinary language untampered with.

Let me briefly recapitulate what was said in section 2.1. That section started out with the claim that the project of explicating symbolic understanding is best served by relying on a conception of agency which does not hinge on the notion of intending. Instead, I argued, we should conceptualize agency in terms of purposive behavior. I then expounded on Short’s theory of purposiveness, according to which an organism’s behavior is purposeful just in case that organism is disposed, by virtue of properties selected for, to flexibly modify its behavior in such a manner that a certain type of outcome thereby occurs (or would occur) more or less consistently. How does this theory bear on the concept of use? For starters, use implies purposiveness: if you use something, this entails you act for a purpose. Use is thus purposive behavior. The verb ‘to use’, furthermore, is transitive: it invariably comes with a grammatical object. To use something is therefore to behave purposefully with regard to something else. When we say that something, or someone, is being used, we often refer to the purpose for which the usage occurs by means of such words and phrases as ‘for’ or ‘(in order) to’. I use my keyboard for writing; a teacher uses examples in order to bring her point across; Hitler used the Gleiwitz incident to justify the invasion of Poland. Purposes can also remain implicit, for instance when one’s aims can be deduced from the context (e.g. “Do you need a spoon? – No, I’ll use a fork.”). But regardless of whether aims are explicitly mentioned or not, an item or event that is used figures as part of an attempt to achieve a purpose. More specifically, it figures as a *means* to a purpose: as something which is (deemed) necessary, sufficient, auxiliary or otherwise contributory to attaining the purpose. So, in sum, when some *X* is used by an agent, that agent uses *X* as a means to achieving some purpose.

The above has merely clarified this section’s explicandum: the way in which the word ‘use’ functions in common parlance. What does my explication of this explicandum consist in? It consists in the replacement of the word ‘use’ by the word ‘selection’. I say that to use something is to select it as a means to a purpose. Further

### 3 Symbolic understanding

on in this section, I shall motivate this statement. But in order to explain how it is to be understood, I first need to say a bit more about what selection is. A useful approach is to first focus upon a type of selection which has already been intensively theorized about: natural selection. As Sober (1984, p. 97-102) explains, natural selection involves *selection-of* and *selection-for*. Assuming a gene-centered view of biological evolution, selection is *of* genes, *for* certain phenotypical characteristics those genes give rise to. Through bestowing a comparative reproductive advantage upon the organism, a certain phenotypical characteristic, along with its associated genetic substrate, comes to predominate in a population at the expense of other phenotype-genotype combinations. At the heart of this selective process is a reduction of variation: a development from a situation in which there is a certain heterogeneity in genotypes within a population, to one in which the distribution of genotypes is at least slightly more homogenous. The prior heterogeneity results from genetic mutations that have occurred at some point in evolutionary history; the posterior homogeneity is due to the fact that environmental conditions differentially affect the reproductive capacities of individual organisms, depending on how those environmental conditions act upon, and interact with, their phenotypes.

Selection, understood simply as variation-reduction, is also encountered outside the sphere of biological evolution. For instance, we can discern a trajectory from relative heterogeneity to relative homogeneity in the way in which a gust of wind ‘acts’ on a collection of chicken feathers strewn around a field in the molting season. I assume that prior to the occurrence of the gust, the feathers fall within a certain range of weights: some are larger and therefore heavier than others. As a result of the collection being affected by the force of the wind, this range becomes more restricted: the feathers that remain in place will, perhaps with some exceptions, be found to exceed a certain weight. The wind thus operates as a selection device: it brings about a reduction in variation through differentially retaining items based on certain properties those items have.<sup>76</sup> The wind, in other words, can be said to select *for* feathers of a certain type. In order to distinguish between this sort of example and natural selection, let us refer to the kind of selection exemplified in the feather case as *mechanical selection*. I explained previously that natural selection amounts to variation-reduction in the distribution genotypes, which occurs by virtue of genotype-determined differences in the reproductive capacities of individuals within a certain population. The concept of mechanical selection pertains to variation-reduction in the distribution of other kinds of types or properties; variation-reduction which does not occur through differential reproductive success, but instead simply by virtue of how items with different types and properties are affected by environmental influences in different ways. Natural selection, as we know, is not purposeful. Neither is mechanical selection. But there is a further, third form of

---

<sup>76</sup> I thank dr. Short for alerting me to the analogy between natural selection and examples such as these.

selection which *is* purpose-directed. This is the selection that is effected by human beings, or, more generally, agents. I thus call it *agential* selection.

How does agential selection work? Take our collection of chicken feathers again, but suppose now that instead of the wind, it is your daughter who has taken it upon her to sort the feathers into two subcollections according to their weights. After a time-consuming and laborious process she ends up with two heaps, one of which contains comparatively small and light feathers, and the other relatively large and heavy ones. Like the wind, your daughter has acted as a selection device: through sorting the initial collection of feathers, she has rearranged it in such a way that both of the resulting subcollections are more homogenous than the original. But unlike the ‘behavior’ of the wind, your daughter’s mysterious rearrangement project was a – supposedly – purposeful endeavor. Her sorting was aimed at some goal, while the wind’s was not. The distinction between your daughter and the wind, when understood in terms of the account of purposeful behavior presented in chapter 2, resides in the fact that the latter lacks a disposition that the former possesses: your daughter’s sorting behavior was liable to flexible modification in the sense that failure, had it occurred, would have been responded to with correction in at least some near-by possible worlds. Admittedly, this disposition to correct mistakes may not have been very robust. For perhaps your daughter would have given up on her project straightaway in case she had inadvertently misallocated a number of feathers in a row. But regardless of your daughter’s (lack of) persistence, her disposition for self-correction will in any case have been more robust than, for instance, the disposition of the atmosphere to produce another gust of wind upon having failed to relocate a feather.

We can now draw two intermediate conclusions with regard to the explication of usage. First of all, the fact that selection is variety-reduction entails that use, by stipulation, also consists in variety-reduction. Secondly, we can infer that usage amounts to agential selection. For as we saw earlier in this section, usage is necessarily purposeful: whenever something is used, it is used as a means to achieving a purpose. Per our explicatum, this means that the selection in which usage consists must be selection that is effected through purposeful behavior, and hence by agents. Some further elucidation of these two intermediate conclusions may be helpful. First, in order to better understand the idea that use is variation-reduction, let us consider an everyday example, such as your use of a fork for the consumption of food. Suppose that, prior to your act of food consumption, there is a certain assortment of cutlery at your disposal: chop sticks, spoons, knives – and forks. Similar to your daughter’s sorting of the feathers, your taking the fork from this assortment boils down to sorting the initial set of items (chop sticks, spoons, knives, forks) into two subsets (one containing the fork, the other containing the rest), according to whether the items actually figure in your act of food consumption. This development amounts to variation-reduction, since both of the resulting subsets harbor less variation than the original. As to the second claim, the following can be



### 3 Symbolic understanding

said. When one uses something, one selects it, in one's capacity as an agent, *for* a certain property it possesses (or for a type to which it belongs), and *as* a means to attaining a purpose. In the case of fork-usage, the fork's role as a means to consuming food consists in its being a vehicle for moving chunks of somewhat soft (but not overly soft) food. And the properties the fork is selected for relate to the way in which it functions as a means. So, in light of the fork's role as a vehicle for food-transportation, selected-for properties could be the fork's pointiness, its easy maneuverability and its tendency to retain food items whilst being moved.

At this point, an objection might be raised to the effect that the proposed explicatum of usage does not satisfy Carnap's similarity condition. According to this objection, the concept of selection does not adequately capture our everyday sense of usage: using something as a means to a purpose seems to go beyond mere selection. After all, usage is intuitively associated with such things as the exercise of a practical skill, active engagement with the object of use, or more generally a certain complexity in behavior – connotations which are apparently absent from the notion of selection (understood as variation-reduction). If anything, selection would seem to *precede* use, rather than constitute it. Now, to label something as an explicatum arguably discharges one of the duty to defend one's proposal against all such objections, provided that the explicatum successfully serves an ulterior purpose which compensates for any loss in intuitive plausibility. But in case an explicatum ostensibly changes the extension of an otherwise unproblematic everyday term significantly, there is nonetheless an onus on the proponent to motivate why his proposal is to be preferred over the default option of leaving the term untouched. I shall therefore reply to the above objection, and try to show that the concept of selection can really capture our common understanding of usage.

According to the objection just raised, selection is what happens before something is used: one *first* picks something, and *then* applies one's 'magic' to it. This idea, however, loses its appeal once we scrutinize more carefully what is actually involved in using something. Such careful scrutiny bears out that ordinary talk of usage typically covers a set of more specific instances of usage. To see this, let us proceed by means of a simple example. If a sailor intends to moor his boat alongside the quay, he might decide to use a bollard as a means to this purpose. But ordinary language, as so often, is not perfectly specific: to say that someone uses a bollard would usually imply that, apart from the bollard itself, there are also other things which are used in connection to, or as part of, using the bollard. For instance, in mooring his boat (and hence in using the bollard), the sailor might use a rope and a reliable knot that is easy to untie, which in itself implicates use at yet another level: that of one's hands and arms. And we could then even go so far as to say that using one's hands and arms further entails the use of certain muscles. Once a 'chain of usage' is untangled in this way, the magic unravels: selection only *seems* to go precede use, because talk of usage is unspecific in a way that talk of selection is not. That is, saying that someone uses a bollard standardly implicates various other forms

of usage, but to say that someone selects a bollard does not obviously carry a similar implication. To make up for any apparent intuitive deficits of the explicatum, it has to come with the proviso that use is often implicitly nested.

Thus far, I have talked only about usage in general, without tailoring the proposed definition to our target concept: *symbol-usage*. What can be said about using a symbol as a means to a purpose? First off, if using symbols amounts to selecting them as means for purposes, then selecting symbols must amount to some sort of variation-reduction being effected by an agent. It is not immediately evident what kind of variation-reduction this could be, however. If a teacher utters “John, pay attention!”, that utterance is used: it is selected as a means for some communicative purpose. But what variation is thereby reduced? Surely the teacher is not picking the utterance from a number of actual alternatives, as one does in taking a fork from the cutlery tray. By the looks of it, no variation is reduced in the teacher’s reprimand. Or is there? Symbol-usage being purposeful, every instance of it is subject to potential correction on the part of the agent. This means that there will almost invariably be a set of possible other symbols an agent would use in case his initial effort in using a symbol would fail (e.g. “John, are you listening?”; “Put away your phone, John!”; “John, stop ruining my day...”). A native language-user might seldom be required to start varying her behavior upon having failed to make herself clear, but the fact that she *would* vary her behavior if her communicative acts *were* to be unsuccessful, attests to the fact that variation-reduction is nonetheless central to her linguistic proficiency. For the presence of her self-corrective disposition indicates that in speaking fluently, she reduces variation: namely the variation she is in principle capable of manifesting. Using a symbol thus amounts to selecting it – not from a repertory of actual items, but from a range of possible alternatives.

One further issue with the idea that symbol-usage is symbol-selection needs to be cleared away. This issue concerns the question what, exactly, the unit of selection is in case of symbol-usage. Is what is selected abstract (a symbol) or concrete (a replicative sign)? One reason for preferring the latter answer is that thinking of symbol-usage in terms of replica-selection seems to better account for what actually goes on. After all, when a teacher utters “John, pay attention!”, that utterance is a replicative sign: it is a particular which can justifiably be interpreted, on the basis of a rule of interpretation, as signifying something (e.g. an imperative directed at some nearby John). So perhaps we should say that symbol-usage ultimately takes replicative signs, and not symbols, as its objects. But this, I would argue, is too quick. Selection, as I have argued, consists in variation-reduction. Has the teacher, in uttering “John, pay attention!”, reduced a range of possible alternatives to the single replica he produces? This, it seems to me, would be to overstate the reduction. After all, the teacher might have used other tokens of the same type of expression that would have suited his communicative purposes equally well. He could have shouted marginally louder, or at a slightly higher pitch. And while it is inevitable that only one particular replica ends up being produced, it is but incidental to the teacher’s act

### 3 Symbolic understanding

of selection which one it is. The reduction, then, is not so much towards one replicative sign, but rather to a set of possible replicative signs that go with the same rule of interpretation. And this implies that variation-reduction is to a symbol, rather than to just one of its replicas. That being said, not all replicas of a symbol may suit some communicative purpose; whispering “John, pay attention!” from afar, will probably not suffice for keeping John focused. So depending on the context, the reduction is likely to be to some proper subset of the entire collection of similar replicas. Still, I would maintain that it is admissible to continue speaking of the selection of symbols simpliciter. In calling someone’s name, sending a text message, or firing a starting pistol, we really select symbols, not replicative signs.<sup>77</sup>

Finally, I should comment on how the proposed explicatum of symbol-usage is meant to enhance the internal coherence of my account. This, after all, is what the explication of the concept of symbol-usage was meant for in the first place. The idea that symbol-usage is symbol-selection can be seen to fit into a broader picture by considering the fact that *all* purposive behavior involves selection. For any bodily movement involved in purposive behavior, insofar as it is liable to self-correction upon failure, forms part of a set of movements from which it is, by that token, selected. Purposiveness presupposes reduction of possible variation, which in turn entails selection. Through defining usage in terms of selection, the notion of usage is tied to the account of purposive behavior presented in chapter 2. For if usage is selection, usage is part and parcel – indeed, the essence – of purposive behavior. In behaving purposefully, we use something – be it an external object or a movement of our own making. So conceived, symbol-usage is nothing but a specific type of purposeful behavior; namely a type in which a symbol figures as a target of agential selection.

#### **3.2 How to use symbols properly**

A symbol, like any other item, may serve a variety of purposes. A speed limit sign can be used in order to warn drivers that they are not to accelerate their vehicles beyond the indicated limit. But that same speed limit sign, taken out of its usual environment, comes to serve a rather different purpose when used to cover up a hole in the floor. Both cases are instances of selection of a means for a purpose. But not all possible purposes of selection concern us equally here: within the purview of explicating symbolic understanding in terms of competent symbol-usage, we are interested merely in those purposes relative to which a symbol is used *as* a symbol. It is such *proper symbol-usage*, after all, which our communicative practices revolve around. Having symbolic understanding thus minimally requires having a disposition for proper symbol-usage. In the first part of this section, I argue that the

---

<sup>77</sup> Something similar can be said about the use of forks: in taking a fork from the cutlery tray, one evidently picks just one particular object, but what is selected is a type, or perhaps a more specific variant of that type (a particular kind of fork).

proper usage of a symbol consists in that symbol's being selected in order to comply with its governing rule of interpretation. Stated differently: to use a symbol properly is to select it for the signficatory purpose associated with that symbol. In the second part, I respond to some concerns that may be raised against the idea that symbolic understanding consists exclusively in behavioral dispositions for proper symbol-selection.

It is useful at this stage to recapitulate some of the main findings from chapter 2. In that chapter, I relied on Short's interpretative work to outline the basics of Peirce's semiotic theory, with some minor (terminological) adjustments. The upshot of that theory is that signification is a triadic relation: signs are elements within a relation that obtains between a sign, an object and (an aspect of) an interpreting response. It was noted that there are subjective and objective signs. An item is a subjective sign when it is *interpreted* as a sign of some object on account of the actual, past, apparent or supposed existence of a prior relation between sign and object. And an item is an objective sign insofar as it is *interpretable* as a sign based on a prior relation that is, indeed, actual. Now, as regards symbols, I proposed that a symbol is first of all a special kind of type: namely a type whose tokens signify. What is more, we found also that symbols are special types of a peculiar breed: their tokens signify by virtue of a purpose for signifying. It was then argued that such tokens, called replicative signs, are governed by rules of interpretation. On the basis of these ideas, the following definition of 'symbol' was formulated:

*Symbol*

A type-sign  $T_{\Sigma}$  is a symbol,  $\Phi$ , if and only if for all token-signs  $\Sigma_i$  belonging to  $T_{\Sigma}$ , some object  $O$ , and some rule of interpretation  $I$ ,  $\Sigma_i$  is a replicative sign ( $\varphi$ ) of  $O$  that has  $I$  as its basis.

Simply stated, symbols are sets of similar replicative signs. Or, somewhat more precisely, they are sets of replicative signs that are governed by a common rule of interpretation.

I concluded in the previous section that the use of a symbol boils down to its being selected as a means for some purpose. This brings us to the main question of this inquiry: what does it mean to say that a symbol is used *competently*? In sections 3.3 and 3.4, I will work my way towards an answer to that question. As a first step, this section will address a further prefatory question: how must a symbol be used by an agent in order for that agent to be at least liable to an ascription of symbolic understanding? As indicated in the introduction to this section, having understanding of a symbol minimally requires that one is disposed to use the symbol *as* a symbol. Whether this criterion is met depends on what purpose is served with a particular instance of usage. A road sign is a replicative sign and hence an instance of a symbol, but it is not used as such when selected as a means for covering up a hole in the floor. After all, it is merely qua physical item of a particular shape and size, and not qua

### 3 Symbolic understanding

symbol, that the road sign serves to cover up the hole. Likewise, we do not use the word ‘cheese’ as that word when we utter it upon a photographer’s request: it is then merely a device for making us smile.<sup>78</sup> Thus, in order for an instance of symbol-selection to possibly qualify as competent, the purpose for which such selection occurs must be germane to the symbol. In other words, to use a symbol qua symbol is to select it for a purpose that corresponds to the purpose of the symbol itself. But what, exactly, is the purpose of a symbol?

Briefly put: a symbol’s purpose consists in the type of outcome to which compliance with its associated rule of interpretation tends to give rise. This was already touched upon in the previous chapter, but I shall belabor the point further here. Recall from section 2.5 that a rule of interpretation is a conditional imperative which says: ‘If a token of *X*’s type is encountered, interpret it as a sign of *O*, or of things of *O*’s type’. Cast in terms of the definitions of signhood from section 2.3: a rule of interpretation is a prior relation which justifies a certain type of response to the instances to which it applies. That is, a rule’s applying to *X* makes it so that, relative to some purpose, *X* can legitimately be interpreted as a sign of *O* (or of *O*’s type). By ‘justified’ or ‘legitimate’ is meant that an interpreting response so qualified contributes at least somewhat to achieving the purpose it is produced for.

Relative to what sort of purpose does a rule of interpretation operate? As I argued in section 2.5, such a purpose must be of a kind that matches the type of outcome to which compliance with the rule (typically) leads. It is perhaps helpful to complement this abstract statement with a concrete example. If you say ‘I am Chris’, expected consequences of producing this utterance are that others subsequently address you as ‘Chris’, use that name to refer to you when talking to mutual acquaintances, and expect you to reliably respond in certain ways upon being called ‘Chris’. Such are the consequences of introducing yourself. And it is relative to those (and related) consequences figuring as purposes of possible interpreting responses that the utterance ‘I am Chris’ constitutes a sign of the utterer declaring her name to be Chris. To be sure, such purposes usually operate below the threshold of our conscious awareness: we are unknowingly complicit in perpetuating the linguistic practices that scaffold our lives. But those purposes nonetheless govern our interactions, given that

---

<sup>78</sup> To clear away possible confusion: while the photographer’s request to say ‘cheese’ is a replica of a symbol (namely an imperative of sorts), our gleefully saying ‘cheese’ in response to that request is not. For that response is in principle not meant to signify anything, whether to the photographer or to anyone else. It is just meant to forge a smile, more or less mechanically. But I say ‘in principle’, because in practice it will often be the case that the mere pronunciation of the word ‘cheese’ is not what actually causes us to smile. Rather, it seems plausible to say that we smile of our own accord, and that our saying ‘cheese’ serves instead to signal our willingness to cooperate, and perhaps a general sense of joyfulness, to the photographer. On such a reconstruction, admittedly, ‘cheese’ does function as a symbol.

they buttress the interpretative rules which make interaction possible in the first place.<sup>79</sup>

To reiterate: the proper usage of a symbol consists in its being selected for the purpose the symbol has. And the purpose the symbol has is the one relative to which its tokens are interpretable as replicative signs of a certain object, or type of object. Thus, a speed limit sign is used properly when selected as a means for ensuring compliance with the speed limit it indicates. That, or some similar aim, is the speed limit sign's purpose, since it is that purpose (or a similar one) which accounts for its interpretative rule's being operative. To illustrate this in more elaborate fashion: suppose that you drive on a country lane and encounter, fitted on top of a steel post, a circular object with a red-colored edge which says '60'. Being ever dutiful, you quickly gauge the speedometer of your car and gently release your foot from the gas pedal until the speedometer says '60'. What have you just done? Among other things, you have committed an act of interpretation: through bringing the speed of your vehicle down, you have interpreted the circular, red-edged object as a speed limit sign. Your purpose in doing so was perhaps to avoid unwelcome financial consequences of maintaining your initial speed, or, more loftier, to ensure a safe and efficient flow of traffic. Those are the kinds of purpose which justify the circular object's being interpreted as a speed limit sign on the basis of an interpretive rule, because it is relative to such purposes that complying with the sign's rule of interpretation makes sense. Makes sense, that is, provided that the sign was used properly by whoever put it there. If it turns out that the road sign was planted there merely in order to allow for cyclists to park their bikes against the sign's metal post, the purposes that would ordinarily justify the object's interpretation as a speed limit sign would now no longer do so.<sup>80</sup>

---

<sup>79</sup> This might be taken to suggest that, despite talk of 'rules of interpretation', my view ultimately reduces to what Paul Horwich (1998, p. 43) calls a "use theory of meaning", akin to the position allegedly adhered to by the later Wittgenstein. The controversial implication of such a theory is that the meanings of linguistic expressions (and, by extension, other symbolic entities) are not constituted by norms. Although this is not the place to tread into these matters in detail, I do not think that the account of meaning entailed by the Peircean framework employed here boils down to a complete denial of normativism about meaning. Rather, it seems to presuppose a kind of 'dual root-theory' of meaning, where norm and use are mutually constitutive of each other and thereby jointly constitutive of meaning. The aspect of mutuality resides in the following. On the one hand, Peirce's account entails that rule-governance forms the possibility condition for the interpretation of essentially arbitrary signs. Without a rule of interpretation, a replicative sign is really just a random object (or sound, or movement). So it would seem we need interpretative rules before symbolic practices can take shape. Yet on the other hand, the content of a rule is established only through actual use: it is only by virtue of actual usage that there is such a thing as a 'typical consequence' to which an interpretative purpose can be matched. From that angle, use would appear to be prior to norms. At present, I see no way out of this paradoxical predicament.

<sup>80</sup> One further case to consider is the scenario in which someone appears to be using a symbol properly, but lacks the authority needed for actually imbuing the item with the intended function. Think, for instance, of the frustrated road-side homeowner who somehow manages

### 3 Symbolic understanding

Before proceeding with the second part of this section, I should like to briefly discuss one further issue surrounding the distinction between symbol-usage simpliciter and proper symbol-usage. This concerns the question whether a replica of a symbol used improperly is still a replica of that symbol. For instance, is the road sign, qua cycle-parking attribute, still a replica of the speed limit symbol it ostensibly instantiates? If we go by the Peircean definitions, it cannot be, since its being interpreted as a replicative sign is no longer justified by the rule of interpretation that would otherwise have governed it. The implication of this is that improper use of a symbol is, in fact, not use *of* a symbol at all. In light of this, speaking of ‘proper usage’ would seem to be redundant: use of a symbol is, by definition, proper. Nevertheless, I have chosen to stick with the redundant phraseology, as it accords well with the way in which labels tend to ‘stick’ to items more or less independently of the purposes they serve. For instance, we might still say “why have you used that road sign as a cycle-parking attribute?” Such linguistic habits attest to the fact that we implicitly continue to recognize the original function of replicative signs (long) after they have ceased to function as such.<sup>81</sup>

We now have a somewhat more precise idea of what using a symbol qua symbol amounts to. At this point, we may want to reflect some more on how the explicatum, around which the first contours have now been drawn, transforms our intuitive conception of symbolic understanding. By centering around the notion of usage, the explicatum retains symbolic understanding’s behavioral dimension, but jettisons its cognitive aspect. Here is why the explicatum can indeed spare no room for the latter. As explained in the previous section, the object of symbol-selection is the symbol as

---

to get hold of a genuine speed limit sign, and places it at his preferred location in order to foster compliance with the indicated speed limit. Has he used the speed limit sign properly? In light of the aim speed limit signs ordinarily have, it seems he has. Then again, the homeowner’s traffic sign does not function in quite the same way that normal ones do. After all, drivers who fail to comply do not commit a punishable speeding offence. I think examples such as these can be accounted for by specifying in more detail the purpose of a symbol. The purpose of normally functioning speed limit signs consists in (1) fostering compliance with a speed limit, through (2) indicating, and being grounded upon, an administratively instituted rule which is legally enforceable. Our protagonist, however persistent his efforts, does not use the symbol for this more specific purpose. For while condition (1) is met, (2) is not: the homeowner’s purpose is to foster compliance with a rule of his own making, to violations of which no administrative sanctions apply. And even in case his initiative were to concur with an administrative rule already in effect, his use of the speed limit sign would still not count as proper. This is because it would still fall short of satisfying condition (2). That is, the intended fostering of compliance would *accord with* an administratively instituted rule, but not be *grounded upon* it. More generally, speed limit signs belong to a category symbols which can be used properly only by certain privileged agents, by virtue of having purposes that can only be pursued by agents who possess the required authority.

<sup>81</sup> How is it then, that we recognize a replicative sign’s former role? Such recognition, I take it, can be regarded as an observation of similarity: we see the object closely resembles other items that we habitually interpret as replicas of a symbol, and on that basis (i.e. on a prior relation of similarity) come to interpret the object as a sign of those replicas.

such and hence an abstract type (or perhaps some proper subset of its associated set of tokens), but what is produced through selection is inevitably a replica. In currently writing these words, I am selecting the types those words are for certain communicative purposes I have. Yet the effect of that selection, once committed to print, is an actual heap of carbon on a piece of paper. A writer selects types, but types tokens. This, in fact, goes for any kind of selection: selection, in order to be actual, must involve the production of a particular instance of the type that is selected. This feature of selection undermines the idea that symbolic understanding, when couched in terms of symbol-usage, can be purely internal. For it is implausible to say that we produce an actual replica of a symbol in processing a symbol cognitively.<sup>82</sup>

I have argued earlier that the cognitive dimension of symbolic understanding can be legitimately sidestepped in light of the function of ascriptions of symbolic understanding: communities simply have no need to exert precise control over the internal lives of their members. Many, however, may be reluctant to give up the intuitive idea that symbolic understanding is partly cognitive. Is there no way to salvage this intuition within the framework currently being developed? In what remains of this section, I will attempt to show how we can restore some of cognition's epistemic relevance whilst staying within the confines of the idea that symbolic understanding consists in symbol-selection. The gist of my proposal is that cognition is ultimately instrumental to, rather than constitutive of, symbolic understanding. That is to say, thought, belief, visualization, or whatever other mental machinery we allow for, plays a role in symbolic understanding insofar as it contributes to acquiring and sustaining behavioral dispositions for selecting symbols properly. Hence, the possession of symbolic understanding may, and indeed very often will, depend on one's cognitive capacities in an indirect sense. To be sure, this does nothing to alter the nature of the explicatum itself. The explicatum still amounts to a behavioristic account of symbolic understanding, in the sense that it says that ascriptions of symbolic understanding apply exclusively to the behavioral dispositions of agents. The intuitive deficit of the explicatum is meant to be compensated for by theoretical gains at the level of its comparative exactness and fruitfulness (see chapter 4). It is therefore mainly to those who are skeptical of my behaviorist outlook *and* who have misgivings about the transactional character of the method of explication, that I present the proposal that I sketched just now, and which I shall work out in a bit more detail below.

As said, we cannot select symbols in our minds. I maintain, however, that there is something else that is selected when we process symbols internally: cognitive states or processes. This requires some elucidation. Recall from section 2.3 that thoughts (and also feelings) constitute auxiliary means to achieving behavioral

---

<sup>82</sup> One could perhaps argue, on the basis of a particular view of the nature of cognition, that the mind produces images or pictures of replicas. But an image or picture of a replica, however similar to it, is not a replica of a symbol.



### 3 Symbolic understanding

purposes. When we think, we put ourselves in a position to select bodily movements and external objects. For instance, thinking your way through a mathematical problem allows you to manifest the solution you take to be right in speech or writing. In manifesting it, you select symbols as means to making your preferred solution amenable to being responded to with – hopefully – unanimous validation by others. The thinking process that preceded the anticipated outcome of this selection procedure therefore formed an indirect means to realizing that outcome. That is, the symbols were selected as means to achieving a purpose (your solution being applauded, or at least not objected to), but the antecedent thought process was what enabled you to select the symbols in the first place. And just as an act of symbol-selection is liable to flexible modification, a thought process can, and often will, be varied upon failure. If someone were to correctly point out a mistake in your solution, you would, presuming you were sufficiently motivated, try to select different symbols in an effort to finally receive the hoped-for validation. As part of engaging in this effort, your original thought process would be altered as well. Such cognitive flexibility, like the flexibility of overt behavior, implicates selection. For as I explained in the previous section, the existence of a disposition to self-correct entails that there is reduction of possible variation, which in turn entails selection.

The purported cognitive dimension of symbolic understanding can thus be found one level ‘below’ that of symbol-usage. Certain thought processes are instrumental to acts of symbol-selection, which are instrumental to achieving the purposes of symbols. In listening and reading, for instance, we do not select the symbols we hear or read – even if we are indeed proficient listeners and readers. The contents of what we read or listen to are not selected – they are simply presented to us. But we do select the cognitive states or processes that we bring to bear upon those contents. Such selection may serve a variety of purposes, not all of which are pertinent to symbolic understanding. Reading and listening carry epistemic relevance insofar as the internal goings-on in which they consist are geared towards developing or preserving the internal pathways that are necessary for engaging in competent symbol-usage. Such internal goings-on are selection-events, for we are inclined to vary them if it turns out that the outcomes sought are not achieved.

A little illustration may help to make my idea more clear. For example, hearing someone say ‘I am Chris’, may evoke in me a cognitive response. I could think, for instance, ‘Ah, this person’s name is Chris’. Such a response, even though it usually occurs automatically and subconsciously, serves to prepare me for future behavior: through recognizing Chris’ utterance as an introduction of herself, I am able to engage in the proper use of symbols at later stages. It may enable me, for example, to say ‘Hi Chris!’ when meeting her on the street.<sup>83</sup> Of course, semantically

---

<sup>83</sup> A subtle misunderstanding threatens to rear its head here. I do not mean to say that I am able to say ‘Hi Chris!’ by virtue of having remembered Chris’ name – although that is obviously true as well. Rather, the point I wish to emphasize is that I am able to say hi to Chris through having recognized her having introduced herself so.

categorizing Chris' utterance in a certain way does not by itself suffice for producing the greeting 'Hi Chris!': I must also understand what 'hi' means, how it is combined with names, in which contexts it is uttered, et cetera. But having internally responded to Chris' utterance in the appropriate way surely helps foment certain dispositions for selecting symbols properly.<sup>84</sup> In sum, I stand with the idea that using a symbol properly is a matter of being disposed to select that symbol in accordance with its rule of interpretation. And while such a disposition is necessarily behavioral, acquiring and sustaining that behavioral disposition will invariably require having certain cognitive dispositions. It is in this sense that cognition plays an indirect, yet crucial role to symbolic understanding.

Still, one could object this cannot be the entire story. For it seems that cognition's epistemic relevance must ultimately be independent from the behavior it renders possible. For instance, people with expressive aphasia can still comprehend perfectly well the ramblings of other people. To base such a person's level of symbolic understanding purely on her behavioral dispositions would be to say that her mental capacities, however appropriate, are epistemically inert. This would be very much at odds with our pretheoretic ideas about symbolic understanding. As I see it, this worry can be alleviated by extending the range of dispositions for symbol-usage of which propriety can be predicated. Besides actual behavioral dispositions, we can include those behavioral dispositions which an agent's cognitive capacities would have allowed for, had her behavioral abilities not been affected by factors squarely outside of her control. On this picture, a person with expressive aphasia possesses symbolic understanding to the extent that the cognitive competence she actually has would have been matched by a corresponding competence in behavior had she not been aphasic. To be clear, this does not detract from the fact that symbolic understanding is still, without exception, a matter of dispositional behavior. Neither does it impinge upon the claim that the role of cognition in symbolic understanding is ultimately only instrumental. But it does indirectly widen the scope of cognition's epistemic relevance, since apart from actual behavioral dispositions, we now also admit a certain class of *counterfactual behavioral dispositions*.

This may come across as an ad hoc solution that serves purely to rectify an all-too-painful intuitive deficit. However, I shall argue that it actually makes a good deal of sense in view of my aim to capture the function of ascriptions of symbolic understanding. That function, as explained before, is to maintain a stable practice of symbol-usage, so that community members can communicate efficiently. In light of the importance of efficient communication, limitations to agents' dispositions for proper symbol-usage are, generally speaking, detrimental to a community's proper

---

<sup>84</sup> I do not wish to insist that the relevant cognitive machinery necessarily comes in the form of thoughts whose contents can be spelled out by linguistic means. The current section makes no claim with regard to the nature of mental states – it only assumes that there is *something* happening internally, the epistemic role of which consists in its being instrumental to the formation of behavioral dispositions that are constitutive of symbolic understanding.

### 3 Symbolic understanding

functioning. Therefore members of a community have good reason to mutually discourage the possession of dispositions for improper symbol-usage and to encourage the possession of dispositions for proper symbol-usage, *provided*, however, that any shortcomings they observe amongst one another are amenable to the kind of behavioral modifications that ascriptions of inferior symbolic understanding are meant to induce. When, however, shortcomings in symbol-usage are not modifiable in this way, community members should not want to use the concept of symbolic understanding in order to improve each other's symbol-using dispositions. Rather, in that case, the community should want to safeguard the inclusion of a maximum number of individuals whose participation in the community's symbol-using activities is of benefit to others and to the community at large. Countenancing counterfactual behavioral dispositions for proper symbol-usage is a means to realizing this aim. For instance, if people with expressive aphasia are recognized as having a strong symbolic understanding by virtue of their counterfactual symbol-using dispositions, they can continue to be recognized as participants within the community's symbol-using activities, instead of being harshly excluded on the basis of their behavioral limitations.

Let me share some further thoughts on this important issue. Clearly, there are few (if any) among us whose dispositions for proper symbol-usage are, in principle, *completely* optimizable. Our internal comprehension tends to exceed our actual behavioral abilities to at least some unmodifiable extent, if only because our active vocabulary is almost inevitably a proper subset of our passive vocabulary. Among the words or idioms of which we have perfect internal comprehension are invariably some for which we lack dispositions for using them properly, simply because they consistently fail to 'pop up' in our minds at the right occasions. So, going by the above proposal, virtually everyone's symbolic understanding at least partly consists in counterfactual behavioral dispositions. But this invites the question why I am continuing to cling so desperately to an exclusively behaviorist conception of symbolic understanding. Why not acknowledge the direct (instead of the merely instrumental) epistemic role of cognition, rather than go for the makeshift measure of including counterfactual dispositions? The reason why I insist on the latter option is that reserving a direct role for cognition would be at odds with the function of the concept of symbolic understanding. That function, as said before, is to preserve stability in a community's communicative practices. Promoting the possession of cognitive dispositions is simply not what a community needs to do in order to preserve such stability – it is beside the point.

The concept of symbolic understanding, when seen as a means for regulating communicative practices, is intrinsically exclusionary to at least some extent. This is because the norms governing a communicative practice are ultimately determined by those agents who enjoy the privilege of having relatively unlimited behavioral abilities. For to be able, as a member of a community, to communicate, is *ipso facto* to co-determine the norms of the communicative practices in that community. Those

members of the community whose communicative abilities are limited, will from the outset have a weaker legislative position as regards the community's norms of communication. This predicament, in conjunction with the fact that the practice of ascribing symbolic understanding is inherently conservative (geared as it is towards maintaining existent norms), has a tendency to create an environment of persistent exclusion. Since such an environment is detrimental to the proper functioning of a community, there needs to be a means for curbing the excesses of symbolic understanding's exclusionary tendencies. The role of counterfactual behavioral dispositions is to provide for such a means.

To summarize the points just made: overall, the practice of ascribing symbolic understanding sees to the preservation of communicative efficiency, through promoting the possession of certain behavioral dispositions. But since the preservation of communicative efficiency tends to come at the cost of a community's internal cohesion and hence of its proper functioning, there needs to be a mechanism to forestall counterproductive exclusion. At the same time, this mechanism should not be overly generous, as that would greatly diminish the effectiveness of the concept of symbolic understanding as a tool for promoting communicative efficiency. This means that it should not just countenance each and every cognitive disposition for proper symbol-comprehension (whatever such dispositions consist in). For that would make the concept of symbolic understanding insufficiently effective as a tool for promoting communicative efficiency. This is why, subsidiary to actual behavioral dispositions, symbolic understanding must instead encompass a limited class of counterfactual behavioral dispositions – namely those counterfactual dispositions the recognition of which allows a community to treat a maximum number of communicatively underprivileged individuals as full-fledged participants in the community's symbol-using practices.<sup>85</sup>

---

<sup>85</sup> Some residual questions remain. My proposal constrains the range of epistemically relevant counterfactual dispositions to those that one's cognitive capacities would have given rise to, were it not for certain insurmountable behavioral limitations. But where do we draw the line between surmountable and insurmountable behavioral limitations? A further question warrants attention too: how do we circumscribe the notion of a behavioral limitation? Consider, for instance, the use of a dictionary. If consulted wisely, it may greatly enhance our capacities for proper symbol-usage. Hence, with some stretch of the imagination, the absence of a dictionary could well be viewed as a behavioral limitation – one that might, in some cases, indeed qualify as insurmountable (perhaps there is no dictionary around for miles). Does this mean that one's symbolic understanding may encompass those counterfactual dispositions which would have been actual had a dictionary been available? How we answer such a question depends on how exactly we conceive of the function of ascribing symbolic understanding. While the general aim of promoting communicative efficiency seems unassailable, how a community conceives of the specifics of this aim may change over time. For instance, the use of (graphic) calculators in secondary education has in recent years normalized to such a degree that the practice of ascribing understanding increasingly tends to accommodate calculator-based success. Sticking to a somewhat broad-strokes characterization of the function of symbolic understanding ascriptions yields an account of symbolic understanding that is more ecumenical, and hence durable, than an

### 3.3 Between propriety and competence

Agents may reliably perform well, yet fall short of being creditable with competence. This is the case when the results of their actions are in keeping with certain standards for success, but the explanation of their success is somehow not really the right one. This consideration, which applies to performances generally, has a bearing on our current topic too. Using a symbol properly is all good and well, but it does not amount to competent symbol-usage as long as one's proper symbol-usage does not stem from the right 'source'. Hence, the normative baseline for symbolic understanding established in the previous section is really only a baseline: something needs to be added to it to complete the picture we are after. Or so the current section will argue. By way of setting the stage, I will first connect the idea just sketched to an account of knowledge long propagated by Ernest Sosa.<sup>86</sup> One of Sosa's central claims, which was already briefly mentioned in section 1.2's historical overview, is that in order to know there needs to be an explanatory link between the truth of one's belief and the process through which that belief was formed. I will argue that a structurally similar criterion applies to symbolic understanding. Unlike Sosa, however, I do not base this claim upon intuitive judgments, but rather, in keeping with my preferred methodology, upon the function that ascriptions of symbolic understanding serve within a community.

In the present project, I try to steer clear of theorizing about understanding on the basis of arguments and views related to the analysis of knowledge. For as I explained in the first chapter, doing so might lead to an unhappy kind of conceptual contamination, as well as to a replication of problems better left untouched. However, given that understanding and knowledge are undeniably at least somewhat kindred concepts, certain similarities in how we approach them are to be expected. At the most general level, both understanding and knowledge serve to indicate certain kinds of success. And where success is concerned, we are often interested in the source or explanation of that success. That is, as part of assessing whether an agent is praiseworthy for having succeeded, we are usually prompted to look for the cause to which her achievement can be ascribed. Luck, for instance, is usually taken to constitute a normatively inferior cause of success in virtually any endeavor.<sup>87</sup> In

---

approach which tries to tailor our concept of symbolic understanding very specifically to the transient norms of a particular era. I thus regard the residual indeterminacy in my notion of the function of symbolic understanding not as a disadvantage, but as an asset.

<sup>86</sup> I thank Anna Höhl for drawing my attention to the connection.

<sup>87</sup> With the possible exception of gambling, where the situation seems to be reversed. If we applaud the winner, we do so on the presumption that his success was due to luck. If it turns out that he cheated, we retract our praise. But it can be questioned whether the praise assigned to the winning gambler is relevantly similar to praise assigned in other contexts of success. Do we really credit the gambler with having achieved something? Instead of awarding credit, it seems rather that we simply celebrate (or, possibly, outwardly celebrate but inwardly resent) the occurrence of a certain type of event (namely the fortuitous gain of money or something else of value), in more or less the same way that we celebrate someone's birthday.

the case of the analysis of knowledge, the need to exclude luck as a legitimate source of success has given rise to repeated attempts at putting up theoretic bulwarks against ever more refined luck-related counterexamples. Some of these attempts can be categorized into a family of views known as virtue epistemology. By and large, virtue epistemologists share a commitment to the idea that for a true belief to qualify as knowledge, its being true has to be attributable to the exercise of an intellectual virtue (Turri, Alfano & Greco, 2019). Sosa (2007, p. 22-23), who pioneered this kind of view and still stands as one of its most influential defenders, employs the following analogy to illustrate it:

When an archer takes aim and shoots, that shot is assessable in three respects.

First, we can assess whether it succeeds in its aim, in hitting the target. Although we can also assess how *accurate* a shot it is, how close to the bull's-eye, we here put degrees aside, in favor of the on/off question: whether it hits the target or not.

Second, we can assess whether it is *adroit*, whether it manifests skill on the part of the archer. Skill too comes in degrees, but here again we focus on the on/off question: whether it manifests relevant skill or not, whether it is or is not adroit.

A shot can be both accurate and adroit, however, without being a success creditable to its author. Take a shot that in normal conditions would have hit the bull's-eye. The wind may be abnormally strong, and just strong enough to divert the arrow so that, in conditions thereafter normal, it would miss the target altogether. However, shifting winds may next guide it gently to the bull's-eye after all. The shot is then accurate and adroit, but not accurate *because* adroit (not sufficiently). So it is not *apt*, and not creditable to the archer.

An archer's shot is thus a performance that can have the AAA structure: accuracy, adroitness, aptness. So can performances generally, at least those that have an aim, even if the aim is not intentional. A shot succeeds if it is aimed intentionally to hit a target and does so.<sup>88</sup>

Knowledge, according to Sosa, likewise requires the existence of an explanatory link that ties success to competence: *in casu* between a belief's truth and the source of its being true. That is, in order for a subject to be credited with knowing that *p* (i.e. to 'aptly believe that *p*'), the explanation of the subject's belief being true must be that it was produced through the exercise of the subject's intellectual virtues.<sup>89</sup>

---

<sup>88</sup> Italics are mine.

<sup>89</sup> As Turri, Alfano & Greco (2019) explain, virtue epistemologists differ amongst each other with regard to what counts as a knowledge-procuring intellectual virtue. For Sosa, as well as for other so-called *virtue reliabilists*, the relevant intellectual virtues include only 'purely' cognitive faculties, such as perception, memory and reasoning. In this respect, virtue

### 3 Symbolic understanding

The common tenet of various virtue-epistemological accounts of knowledge, here exemplified by Sosa's 'AAA-structure' of competent performance, is relevant also to the present investigation. Of course, we should be mindful of the fact that knowledge and my notion of symbolic understanding come apart with regard to their respective objects of ascription. Symbolic understanding, as understood in this thesis, consists in actual or counterfactual dispositions for symbol-usage, rather than in beliefs. Understanding thus pertains to types of behavior, whereas knowledge pertains to belief.<sup>90</sup> More importantly, I have resolved in chapter 1 not to tailor my explicatum of symbolic understanding to our intuitions – as is of course common in the study of knowledge – but instead to the concept's function. Notwithstanding these differences, my view is that understanding symbols is, like knowledge, an achievement which requires that the accuracy of an agent's (possible) performance is due to appropriate causes. Further on in this section, I will support this view by arguing that certain ways of acquiring dispositions for proper symbol-usage had better be discouraged lest ascriptions of symbolic understanding lose their utility for promoting communicative efficiency. Prior to presenting this argument, however, it is appropriate to go over some everyday examples which bear out the relevance of Sosa's point to the current topic. I should emphasize that the appeals to intuition which are made in the examples serve a heuristic rather an argumentative purpose: they are meant to acquaint the reader with, rather than persuade her of, my position.

Consider, for instance, the radiotelegraph operator who has a propensity for quickly producing correct Morse signals. Based purely on the consistency of his performance, we might initially be inclined to deem him a competent symbol-user. But if we were to learn that, incredibly enough, his disposition for proper symbol-usage derived from sheer luck, we would surely retract our praise. In addition to luck, there are other sources of proper symbol-usage that make us refrain from

---

reliabilism stands opposed to *virtue responsibilism*, according to which we also need to take into account such traits as open-mindedness, scrupulousness, and impartiality.

<sup>90</sup> Apart from the obvious difference between symbolic understanding being about behavior vs. knowledge being about belief, there is nonetheless a potential commonality between the two. Whether the commonality is deemed extant depends on whether one thinks of beliefs as occurrent (or 'episodic'), or as dispositional. Campbell (1967, p. 206) argues that occurrent belief is the more fundamental notion, given that it would be impossible to specify the nature of a dispositional belief without making reference to occurrent belief: "Among the 'tendencies to react' in terms of which dispositional belief is defined we must surely include, in every case of dispositional belief, the tendency to react to some of the relevant situations with *episodic* belief." But determining which of the two has conceptual priority over the other, is independent from answering the question which of the two is of prime epistemological importance. And here it seems consensus is on the side of the dispositional conception of belief. As Rose & Schaffer (2013, p. 23) point out, the dispositional conception is implicated in the widely endorsed thesis that knowledge entails belief: if we were to countenance only episodic belief, we would know far too little. The relevance of all this is that on a dispositional account of belief, my account of symbolic understanding and standard analyses of knowledge would share in the assumption that epistemic success is ascribed of types, and hence of abstracta.

attributing symbol-using competence. Take, for example, the case of a diligent student who has a habit of internalizing his reading materials so thoroughly that he consistently passes his exams purely through consulting his memory. Or think of a parakeet that utters ‘Hello!’ whenever someone enters the room, through having formed a habit for doing so on the basis of mimicry and conditioning by a human. Like luck, the success-sources alluded to here – mere memorization and brute imitation – are, intuitively, epistemically defective sources of proper symbol-usage.

Before I turn to presenting my actual argument for the view that competence requires appropriately grounded propriety, I should discuss a possible qualm one might have with the examples just given. For one could object that neither luck, memorization nor imitation can really be causes of *proper* symbol-usage. After all, it seems that irrespective of what symbol-usage they might give rise to, such usage cannot possibly be directed at the purpose for which the symbol exists. For instance, lucky understanding would appear to amount only to *apparent* propriety: the radio operator who is disposed, by sheer luck, to efficiently converse in Morse code happens to produce replicas of the right symbols in the appropriate circumstances, but not for the right purposes. Something similar would seem to go for the astute memorizer: he spits out replicas from the top of his head, but without a concomitant aim to achieve the purposes those replicas ordinarily are for – he just aims to replicate them. And again, imitation by a parakeet, even if seemingly indicative of linguistic competence, appears to have to occur relative to a purpose unrelated to the purpose of the symbol being used. It might be conjectured that the bird utters the sound in order to receive food or attention, rather than for realizing those outcomes that correspond to the purpose of the symbol. Such considerations purport to show that what I previously construed as epistemically defective sources of proper symbol-usage, are not really sources of proper symbol-usage at all. One could thus complain that the examples fail to illustrate that there is any gap between propriety and competence which needs to be bridged.

I maintain that this conclusion should be resisted, however. The examples just offered readily lure us into thinking of them as being about particular occurrences of symbol-usage, rather than about dispositions. And I agree that it is indeed somewhat difficult to see how a particular occurrence of symbol-usage that is due to luck, memorization, or imitation, could still be proper in the sense of the symbol being selected for the purpose for which it exists. But we should continue to bear in mind that we are concerned here not with *occurrent* proper symbol-usage, but with *dispositional* proper symbol-usage. When we ask whether someone’s proper symbol-usage amounts to competence, we want to know how it is that someone is disposed to engage in proper symbol-usage. Answering that question is a matter of identifying the properties or processes responsible for the agent’s having the disposition. And it seems to me that so conceived, there is room for soundly, as well as unsoundly anchored proper symbol-usage. One may be disposed to select a symbol in accordance with its purpose, but the explanation of the presence of this



### 3 Symbolic understanding

disposition may still be the wrong one. A habit for doing something the right way can stem from a spurious process of habit-acquisition. In order for symbols to be used competently, one's disposition for proper symbol-usage must have been acquired by appropriate means, and luck, memorization and imitation intuitively do not qualify as such.

A mere intuitive argument is not enough, however. Given my chosen methodology, I must show that dispositions for proper symbol-usage grounded in luck, memorization and imitation should not qualify for ascriptions of symbolic understanding because such ascriptions would fail to contribute sufficiently to the goal of promoting communicative efficiency. Let me start my argument with a simple observation. It is clear that the dispositions for proper symbol-usage that luck, memorization and imitation can procure (if at all) are unavoidably rather limited. Luck normally assist us only for brief bouts – rendering the production of stable dispositions all but unfeasible. Somewhat analogously, memorizing the contents of a text by heart enables us to reproduce those contents faithfully, but it usually does not allow us to do much else besides. And the parakeet that consistently greets incoming visitors is likely to require very specific situational cues in order for his disposition to materialize, such that that disposition fails to generalize to marginally different contexts in which an ordinary hello-sayer might reasonably be expected to greet incoming people too.

The fact that luck, memorization and imitation are not really suitable means for acquiring comprehensive dispositions for proper symbol-usage, renders them ill-suited for functioning as cornerstones of our communicative practices. That is, even if there exist agents with luck-, memorization-, or imitation-based dispositions which are as good as one could wish for, a community cannot expect those dispositions to be readily acquirable by the majority of its members. Stated differently: although luck, memory and imitation might work for some, these are not the pillars upon which to erect an inclusive and stably functioning communicative practice. It is not difficult to see why. Luck, being luck, is unreliable. Memory, although in principle capable of being reliable, requires too much of our cognitive faculties in order for general and robust symbol-using dispositions to arise from it in sufficient numbers. And imitation, while potentially reliable too, is likewise impracticable since it demands the frequent and predictable presence of that-which-is-imitated.

The function of applying the concept of symbolic understanding is to promote efficient communication through encouraging the possession of dispositions for proper symbol-usage. But in light of the above, not all ways of acquiring such dispositions are born equal. Were the concept of symbolic understanding to be insensitive to the manner of disposition-acquisition, it would not be specific enough as a means for fostering efficient communication. Such an overly permissive concept of symbolic understanding would not just be *somewhat* less effective than a more restrictive one. It would be *far* less effective. For it is important to realize that how a community chooses to regulate its communication impacts upon many things,

among which the organization of its educational practices. After all, how we choose to transfer understanding of foreign languages and other symbol systems in educational institutions is predicated on what we take understanding to consist in. If symbolic understanding were insensitive to the manner of disposition-formation, ascriptions of symbolic understanding by teachers could not be used to nudge students towards using fruitful and efficient methods for language acquisition. Granted, such ascriptions would still encourage the possession of dispositions for proper symbol-usage, but they would fail to appropriately reward those agents who develop their abilities in the most efficient ways. In Darwinian terms: a concept of symbolic understanding that did not see to the manner of disposition-acquisition could not be used to apply the ‘selection pressure’ that is needed in order for a community to efficiently instill the required symbol-using dispositions in its members.

Based on the foregoing, I stand by the claim that symbolic understanding is a success term which is predicated of certain behavioral dispositions that are acquired in the right ways. This makes for a superficial analogy between my account of symbolic understanding and virtue-epistemological theories of knowledge such as Sosa’s. Superficial, because I have argued for my position via a different route than Sosa: namely by showing that a concept of symbolic understanding must be sensitive to differences in how agents acquire dispositions for proper symbol-usage, lest the concept falls short of fulfilling its function. The question that now remains is how the gap between propriety and competence is to be bridged. For I have shown that symbolic understanding, understood as competent symbol-usage, goes beyond mere proper symbol-usage. But I have not yet shown in what way it does so. Which ingredient needs to be added to propriety in order to turn it into competence? This question will be addressed in the next section.

#### **3.4 How to use symbols competently**

We now arrive at a crucial point in this investigation: if not luck, memorization or imitation, what *does* constitute an epistemically sound basis for proper symbol-usage? This section will propose an answer to that question. I submit that in order for a disposition for proper symbol-usage to license an ascription of competence, that disposition has to stem from the operation of a learning mechanism. In the first part of the section I will discuss – very concisely – some of the recent literature on learning in cognitive psychology, and draw the cautious conclusion that the learning of symbols is likely to be a multi-faceted affair which resists straightforward unification into one overarching account. Then, I will argue that regardless of which perspective one prefers, there is a crucial assumption without which none of the main theories of learning can do. This has to do with a phenomenon called *selective attention*. As I hope to show, it is only by selectively attending to a limited set of features that we are able to apply a learning mechanism – *any* learning mechanism – to the instances we observe. I conclude this section by condensing the main

### 3 Symbolic understanding

findings of this section into a working definition of competent symbol-usage. Importantly, this definition does not constitute the end product of this chapter, given that it does not spell out what is meant by *better* competent symbol-usage. I will postpone this task to sections 3.6 and 3.7.

As pointed out at the end of the previous section, in order for a community to have an efficient system of communication, it must be the case that sufficiently many of its members can acquire comprehensive dispositions for proper symbol-usage in efficient and readily accessible ways. I think it is safe to say that humans generally do quite well in this regard. This prompts us to look at how, in actuality, dispositions for proper symbol-usage are acquired. Parents alert their children to animals and attach to those animals the names of the species they belong to; high school students become acquainted with mathematical formulae by watching the teacher solve problems to which they apply; astrologically inclined people may acquire the ability to interpret horoscopes by watching other people do so in You-Tube videos. What can be gleaned from such examples? A common aspect that can be discerned is that aspiring symbol-users get a hold on things that are general on the basis of the observation of particulars. The animal pointed to, the introductory problem discussed, or the horoscope interpreted are just singular cases, but they somehow enable us to master the types those singular cases instantiate. How does this work? We already know of at least two mechanisms that do *not* suffice for attaining mastery over types via particulars: mere memorization and imitation. For while these allow for some grasp of types, they only do so in a very limited sense. To retain a memory of a particular animal being pointed to and named in a certain way, is not yet to grasp what category the name generally applies to. It only amounts to memorizing that *that* individual goes by that name. Similarly, imitating someone or something amounts only to replicating a specific occurrence in a rather tightly circumscribed way. As was pointed out in the previous section, this is the reason why mere memorization and imitation must fail to uphold a stable practice of communication: they are simply too demanding and impractical to procure the needed dispositions for proper symbol-usage.

What is it then that allows us to short-circuit the lengthy path from individual instances to types in an efficient yet reliable way? My unremarkable answer is: learning. Below, I try to say a bit more about what learning is by looking into some of the cognitive science literature on this topic. However, since I want to avoid taking positions within scientific debates, and as I strive to make my explication of symbolic understanding as widely applicable as possible, my conclusions will be highly general. So general, in fact, that they are likely to extend beyond humans, thus salvaging the possibility that the concept of symbolic understanding has application outside of the human sphere.

One way in which the scientific literature on learning already caters to my need for generality is that it concerns not merely the learning of symbols, but the learning of all kinds of types. A superficial survey of decades of research in cognitive

psychology reveals two general possibilities regarding the nature of learning types, or, as is the more usual term, *categories*: either we learn categories in a top-down manner (e.g. through applying rules that specify type-membership), or do so bottom-up (e.g. by recognizing similarities between individual instances). The former view, which draws on a theory of categories that can be traced back to Aristotle, finds its canonical expression in Jerome Bruner's *A Study of Thinking* (1956). The latter view has its philosophical roots in Wittgenstein's *Philosophical Investigations* (1953). In the contemporary literature on category learning, the Aristotelian view is usually referred to as the *classical view* (Smith & Medin, 1981). The Wittgensteinian perspective, on the other hand, is present in the so-called *prototype* and *exemplar views*.

According to the classical view, people rely on definite criteria in order to get to grips with the categories that are picked out by certain concepts. That is, they determine whether some item belongs to a particular category by checking whether that item satisfies the necessary and sufficient conditions for membership of that category. With respect to our own subject matter, this would imply that people learn to use a symbol properly by internalizing the rule that governs it, and, more specifically, through becoming familiar with the entry requirements for the set of replicas to which a symbolic rule applies. The prototype and exemplar accounts, by contrast, entail that category-membership is determined via the recognition of resemblance relations among instances of a category. Again, tailored to our present topic, such a view would imply that people learn to use symbols properly by acquiring the ability to associate given items with previously observed replicas of symbols on the basis of similarities between the two.

Before discussing some of the relative merits of these views, we need to briefly zoom in on a subtlety that separates prototype from exemplar accounts. The prototype view, first of all, "assumes that there is a summary representation of the category, called a prototype, which consists of some central tendency of the features of the category members. Classification is determined by similarity to the prototype" (Ross & Makin, 1999, p. 208).<sup>91</sup> Proponents of the exemplar view, on the other hand, "assume that categories consist of a set of exemplars and that the classification of new instances is by their similarity to these stored exemplars" (idem, p. 212). The main point of divergence between the two views can thus be captured by saying that the prototype account, unlike the exemplar theory, posits the existence of an intermediate step of abstraction from observed instances. On the whole, however, both prototype and exemplar theorists are in agreement as to the idea that similarity

---

<sup>91</sup> Kruschke (2005, p. 188) adds that there are other possible ways of concretizing the notion of a prototype: "Another option for the prototype could be an idealized or extreme case that is maximally distinctive from other categories. (...) Alternatively, a prototype could be the most frequent, i.e. modal, instance; or a prototype might consist of a combination of the most frequent or modal features of the instances."

### 3 Symbolic understanding

recognition constitutes the means by which novel instances are categorized. In this sense, they are allied in their opposition to the classical view.

By and large, experimental results have prompted scientists to move away from the classical view, and embrace prototype and exemplar theories.<sup>92</sup> Perhaps the most significant disadvantage of the classical view is that it has difficulty accounting for so-called *typicality effects*: the finding that people's adeptness at producing and recognizing instances of categories tends to vary depending on the item at hand. That is, certain instances of categories are privileged over others in terms of how quickly subjects are able to recognize or produce them and how frequently they cite them. A blackbird is, in a way, more birdlike than a penguin; a beech more treelike than a yew. The classical view, according to which an item either does or does not satisfy the requirements for category membership, has difficulty explaining the occurrence of typicality effects. By contrast, the prototype and exemplar views are able to account for the existence of hierarchies among instances of categories. The prototype view does so on the basis of the fact that instances are bound to differ with respect to their comparative similarity to the prototype. The exemplar view, on the other hand, predicts typicality effects through assuming that "typical instances usually occur more frequently (so there are more exemplars of typical instances) and because typical instances tend to be highly similar to other instances" (idem, p. 214).<sup>93</sup>

---

<sup>92</sup> A primitive, anecdotal version of this view is implicit in the following words of Thomas Kuhn (1977b, p. 305):

Students of physics regularly report that they have read through a chapter of their text, understood it perfectly, but nonetheless had difficulty solving the problems at the end of the chapter. Almost invariably their difficulty is in setting up the appropriate equations, in relating the words and examples given in the text to the particular problems they are asked to solve. Ordinarily, also, those difficulties dissolve in the same way. The student discovers a way to see his problem as like a problem he has already encountered. Once that likeness or analogy has been seen, only manipulative difficulties remain.

A similar observation, Kuhn subsequently notes, can be made with regard to the history of science: many important scientific innovations have resulted from scientists taking their cue from previously established problem solutions and extending those solutions to new cases on the basis of observed resemblances. As Kuhn (idem, p. 307) goes on to argue, similarity relationships, perhaps even more so than rules, are crucial to learning how to recognize and apply types: "Acquiring an arsenal of exemplars, just as much as symbolic generalizations, is integral to the process by which a student gains access to the cognitive achievements of his disciplinary group."

<sup>93</sup> Apart from this empirical argument, there is a deeper reason for thinking that rules, even if we assume that they can assist us in learning, are in some sense secondary to exemplars or prototypes. For a rule, being itself a type, must first be learnt in order for it to be of any use for learning to recognize and apply the type to which it applies. Supposing that rules are learnt by means of higher-order rules is not of much help: for how are these higher-order rules to be learnt? At base, therefore, rule-learning must be effected through the observation of particulars: namely by being confronted with the consequences of compliance and non-compliance. It is ultimately through recognizing patterns of similarity among the

Notwithstanding the scientific consensus regarding the inadequacy of the classical view, more than fifty years of empirical research into category learning has not yielded a conclusive verdict as to whether we should opt for a prototype or an exemplar model of category learning. On the one hand, some experimental results indicate that reliance on prototypes cannot be the full story. Whittlesea (1987), for instance, has shown that equal similarity of items to some prototype does not always run in tandem with people's performance in classification: if one out of a pair of such equally similar items is more similar to an observed instance than the other, people tend to classify it more accurately. On the other hand, exemplar theory might not be wholly satisfying either. Ross & Makin (1999, p. 215) point out that exemplar theorists may be criticized for their outright denial that abstractions play any role in learning. After all, abstractions such as criteria and idealized representations often very much *seem* to aid us in classification, both intuitively and in practice. Is it really so plausible to claim that our reliance on such abstractions is all merely apparent? As Kruschke (2005, p. 190) concludes, we may have to accept that none of the two main theories of category learning taken on its own can exhaustively account for the wide variety of cognitive tasks that fall under category learning. According to a growing number of authors, it seems rather more reasonable to say that humans are capable of utilizing more than one learning mechanism, and that, depending on the nature of the categorization task at hand as well as on contextual factors, distinct learning mechanisms interact and compete with one another in different ways.

Still, there is a sense in which each of the models discussed here – including the classical view – relies on a shared assumption. This concerns a phenomenon called *selective attention*. The phenomenon is so familiar as to easily escape notice. We selectively attend when we examine the soil in a pot so as to check whether our plant requires watering, or when we listen to the blasts from a ship horn in order to determine the skipper's intentions. In such cases, we single out a specific feature, or set of features, and ignore or attenuate others. Potting soil has a certain weight and smell, but for plant maintenance it is mostly its moisture level which concerns us. Likewise, the blasts of a ship horn have both amplitude and frequency, but when it comes to probing the mind of the skipper it is only their number which is of importance. Why does selective attention exist and how does it impinge upon each of the three learning mechanisms discussed? For this, we first of all need to home in on a fundamental aspect of existence. Survival, for any sublunar being, is a matter of dealing prudently and efficiently with scarce goods. Time, space and material as well as immaterial resources are inevitably at least somewhat limited. The items and organisms that enable our survival, however, possess uncountable features. This fact of life necessitates that we continuously disregard a great many aspects of the things

---

consequences of our actions, that we start to grasp the contents of the rules that govern those actions. Nonetheless, it remains possible of course that rules, once learnt, can be aids to category learning. I will therefore continue to recognize the classical view as a candidate theory of category learning.

### 3 Symbolic understanding

we observe, and focus primarily on what is at least somewhat relevant to our continued existence. Only by selectively attending to certain features of objects are we able to recognize and apply the categories that are of immediate or more distant relevance to our survival. As I now hope to show, selective attention is required regardless of which learning mechanism is thought to be involved in acquiring a capacity for category recognition and application. Both exemplar and prototype models, as well as rule-based ones, need to build in a selective attention parameter lest they become implausible.

According to exemplar theorists, first of all, we learn to categorize through making similarity judgments on the basis of stored representations of observed instances. But in doing so, we invariably pay heed to only a limited number of features. That is, we do not judge overall similarity *per se*, but rather similarity with respect to a restricted amount of observable aspects. Take the inscription ‘A’. It is recognized as an instance of the first letter of the Latin alphabet on the basis of specific characteristics of its appearance, such as the size and configuration of the strokes that constitute it. Its geographical coordinates, the depth of its imprint, or the saturation of its color, by contrast, are not taken into consideration when determining the item’s similarity to stored exemplars. Were we to take all such features into account, the comparison of observed instances to stored exemplars would be so cumbersome as to be infeasible.

Selective attention is similarly pertinent to prototype models of category learning. As explained before, prototypes are summary representations of observed instances that serve as intermediaries in category assignment. Such summary representations instantiate a certain set of properties prevalent among observed instances, thereby possibly allowing for a swifter comparison procedure relative to the more demanding process of exemplar-matching. One could argue that prototypes do away with the need for selective attention in observation, given that irrelevant features have already been weeded out. But this is not really the case. A prototype, qua summary representation, still inevitably instantiates properties that are irrelevant to its purported role in learning. For instance, a prototype for the geometrical figure of a square, however minimalistic, has at least some aspects, such as the color and thickness of its contours, which do not really bear upon its function qua prototype. Hence, although it is not unlikely that there is indeed some attenuation of irrelevant features going on in the prototype-formation stage, classification of observed instances must happen on the basis of a limited subset of properties of the prototype.<sup>94</sup> Prototype theorists, like exemplar theorists, therefore have to account for the role of selective attention too.

---

<sup>94</sup> Ross & Makin (1999, p. 212) remark that “the prototype view has never presented a clear learning theory – how is information about instances combined to form a prototype? There appears to be an assumption of a type of associationist learning view, but the learning of prototypes has not been detailed”.

Rule-based theories of learning, finally, cannot pass over selective attention either. The classical view has it that category recognition or application requires that we ‘scan’ the item of interest for the possession of necessary and sufficient features dictated by a rule. But how do we determine which rule, out of an indefinite number of possibilities, we should take into consideration? For this, I take it, we need to selectively attend to a limited set of features of the item. To see this, consider once more the mark ‘A’. According to rule-based models, determining whether this inscription belongs to the category denoted by ‘the first letter of the Latin alphabet’ requires that one checks whether it satisfies the antecedent of a conditional imperative that says ‘if you encounter an item with such-and-such features, interpret it as the first letter of the Latin alphabet’. However, since the mark could in principle be governed by any rule whatsoever, we must have some means of singling out a particular rule for being considered as part of our checking procedure. Such means are provided by our capacity for selective attention: it is through concentrating only on certain aspects of the inscription that we are able to restrict the range of rules that could be of possible relevance. By disregarding such properties as the mark’s geographical location and the depth of its imprint, and by instead concentrating upon its shape and size, we are able to pick out a rule that is of potential relevance in light of the aspects attended to.

Earlier in this section, I submitted that the gap between propriety and competence is bridged by learning. Let us now see in what way the above discussion about the nature of learning bears on our explicatum: in light of what was said, how should we define the notion of competent symbol-usage? For starters, the current state of scientific research regarding category learning does not permit us to be partisan with respect to either of the three main learning mechanisms mentioned. We have to allow for the possibility that the acquisition of dispositions for proper symbol-usage is a highly complicated affair in which different learning mechanisms may be involved to varying extents depending on the context and the nature of the symbol. Consequently, it would seem that we must rest content with the following platitude:

Competent symbol-usage

An agent  $A$  is a competent user of a symbol  $\Phi$  if and only if

- (i)  $A$  has a disposition for using  $\Phi$  properly, and
- (ii) (i) is due to  $A$ ’s having engaged in a learning process pertaining to  $\Phi$ .

But the considerations concerning selective attention occasion a somewhat richer picture. It was shown that, irrespective of the mechanism involved in learning a symbol, any such mechanism can only become operative provided a limited selection of features of an observed instance is attended to. We thus arrive at the following:

Competent symbol-usage (improved version)

An agent  $A$  is a competent user of a symbol  $\Phi$  if and only if



### 3 Symbolic understanding

- (i)  $A$  has a disposition for using  $\Phi$  properly, and
- (ii) (i) is due to  $A$ 's having engaged in a learning process pertaining to  $\Phi$ , which was made possible by  $A$ 's having selectively attended to a limited number of features of one or more replicas of  $\Phi$ .<sup>95</sup>

A question that remains is: what features, exactly? Which aspects should we selectively attend to if our observations are to constitute useful input to a learning process? The answer is: the features that are relevant to the symbolic function of a replicative sign. For example, if we want to properly categorize the mark 'A', we had better attend to the configuration of its strokes, and disregard its geographical location.<sup>96</sup> Only the former aspect, after all, bears on whether and how the mark signifies qua replicative sign. Hence, we can update the latest version of the definition of competent symbol-usage as follows:

*Competent symbol-usage (second improved version)*

An agent  $A$  is a competent user of a symbol  $\Phi$  if and only if

- (i)  $A$  has a disposition for using  $\Phi$  properly, and
- (ii) (i) is due to  $A$ 's having engaged in a learning process pertaining to  $\Phi$ , which was made possible by  $A$ 's having selectively attended to symbolically relevant features of one or more replicas of  $\Phi$ .

As a final tweak to our definition, we should incorporate into it an insight from section 3.2: the idea that symbolic understanding is sometimes attributable on the basis of counterfactual behavioral dispositions. That is, even if an agent lacks some disposition for proper symbol-usage, she may still be credited with symbolic understanding in case she would have possessed that disposition had she not been limited in her behavioral possibilities through circumstances outside of her control. To factor in this exemption clause, the definition should be modified thus:

*Competent symbol-usage (third improved version)*

An agent  $A$  is a competent user of a symbol  $\Phi$  if and only if either

- (i)  $A$  has a disposition for using  $\Phi$  properly, and
- (ii) (i) is due to  $A$ 's having engaged in a learning process pertaining to  $\Phi$ , which was made possible by  $A$ 's having selectively attended to symbolically relevant features of one or more replicas of  $\Phi$ ,

---

<sup>95</sup> Note that whereas condition (i) cites a disposition, condition (ii) makes reference to occurrences. An agent needs to *actually* do things (learn through selectively attending) in order to acquire the *potential* to do other things (use symbols properly).

<sup>96</sup> As I explained in section 2.5 (and as I shall repeat again in 3.5), the relevant features are not limited to purely physical properties: it matters also whether certain physical properties are exploited for a signficatory purpose. In other words, the mark 'A' is not an A unless it is meant to be interpreted as such on the basis of its similarity to other replicas of the letter A.

or

A would have satisfied conditions (i) and (ii) had she not been limited in her behavioral possibilities by circumstances outside of her control.

This, then, is my final word on the nature of competent symbol-usage. It is not, however, my final word on the nature of symbolic understanding. For I still need to gradualize the above definition in order to make good on my earlier resolution to take comparative assessments of symbolic understanding (as opposed to symbolic understanding simpliciter) as the core unit of analysis. This I shall do in sections 3.6 and 3.7. The next section will take a closer look at the phenomenon of selective attention, which, after all, has now come to figure prominently in my explicatum of symbolic understanding. More specifically, I will offer an account of what selective attention consists in, once more on the basis of Peircean conceptual machinery. Furthermore, I shall make a stab at explaining how it is that agents are able to attend selectively. I should note that the claims I am going to put forward, some of which rather speculative, will not be very consequential to the account at large: what follows in sections 3.6 and 3.7 (or in chapter 4, for that matter) can for the most part be understood independently of what is said in section 3.5.

### **3.5 Interlude: the nature of selective attention**

My survey of the literature on category learning revealed that there is no clear consensus on how people learn to recognize and apply categories. I did, however, discern a red thread. Whether learning is rooted in stored exemplars, abstracted prototypes, or in the application of rules; such proposed mechanisms are all assumed to rely on a peculiar operation of the mind: the focusing of attention upon a limited number of features of an item, namely upon those features that are relevant to the item's membership of the category at hand. This insight prompted the inclusion of a selective attention criterion within the definition of competent symbol-usage. But what does that criterion entail, exactly? To answer this question, the current section aims to shed light on what it means to selectively attend to symbolically relevant features of replicas. I will once more bring to bear upon this inquiry the ideas of Charles Sanders Peirce, suggesting that to selectively attend to certain features of an item is to treat those features as so-called icons. Having done so, I shall subsequently clarify under what conditions a feature is relevant to the symbolic function of a replica, and offer a tentative explanation of how it is that we are able to direct our attention to such symbolically relevant features.

Although this section approaches the phenomenon of selective attention from a philosophical standpoint, it is important to briefly discuss some of the scientific insights that have been acquired ever since cognitive psychologists started to recognize the vital importance of selective attention for virtually all cognitive tasks. During the early stages of research into selective attention, the phenomenon was commonly conceived of in terms of a 'bottleneck model' (McLeod, 2018).

### 3 Symbolic understanding

According to Broadbent (1958), for instance, selective attention can be understood as a process in which an undifferentiated array of sensory information enters a so-called ‘sensory buffer’ from which, depending on the physical characteristics of what is observed, only certain stimuli are subsequently allowed to pass the selective filter for further processing. In response to Broadbent’s model, other authors sympathetic to the bottleneck paradigm have objected that empirical evidence suggests that filtering is more likely to occur after instead of before processing a stimulus (Deutsch & Deutsch, 1963), thus provoking a continued debate between proponents of *early-selection* versus *late-selection* theories of attention. More recent work on selective attention has tended to make use of a different model, which emphasizes the role of the innate limits to our perceptual and cognitive capacities. According to the influential *perceptual load theory* of selective attention, for example, the seemingly conflicting implications of early- and late-selection theories can be jointly accounted for in terms of the idea that “the amount of perceptual and cognitive load regulates how effectively stimuli are selectively attended to, determining whether the response will mimic the early versus late selection process” (Bater & Jordan, 2019).<sup>97</sup> Although the perceptual load theory currently enjoys widespread support, some problems persist. As Bater & Jordan (idem) point out, for instance, it is not entirely clear how the notion of perceptual load is to be operationalized. As long as no such operationalization is settled on, interpretations of experimental results may continue to diverge.

---

<sup>97</sup> Murphy, Groeger & Greene (2016, p. 1318) offer the following helpful illustration of how perceptual load theory is deemed to unify the evidence for early- and late-selection theories:

[H]ow does one read this article while ignoring a fly buzzing around the room? How is it that this sentence is selected and the fly is rejected? Early-selection theory dictates that as the fly is irrelevant, it would not be processed. This page would be selected for further attention at an early stage, and nothing about the fly would be processed beyond that point. Late-selection theory suggests that the fly (and other surrounding stimuli) would be processed along with the page. Selective attention would then take place at a later stage in processing, preventing the distractors from affecting behavior. Load theory incorporates aspects of both early and late selection to explain this scenario. If the visual properties of this article incurred high perceptual load (e.g., if the paper was transparent and the words written on the reverse of this page were visible here, demanding increased attention to distinguish these relevant words from the irrelevant distractor words), it is likely that the fly would be filtered out of the reader’s awareness at the perceptual stage and not processed further; early selection occurs in this case because perceptual capacity is exhausted. If, however, the article incurred lower perceptual load (e.g., due to being written on thick, white paper with no translucent properties), the fly would be processed along with the page to a later stage of processing at which the reader must select the page and prevent the fly from interfering with the primary task. The allocation of attention at this late stage is dependent on the available cognitive resources, and it may fail if cognitive load is high. The process of selective attention is, according to load theory, dependent on both external properties (perceptual load) and internal properties (cognitive load).

To maintain an impartial position with respect to science-internal issues I will not side with any of the theories just sketched, but instead draw a connection between selective attention and Peirce's theory of signs, suggesting that the former can be cast in terms of the latter at a very general, scientifically neutral level. I should note that this suggestion only targets the *what* of selective attention; its *how* will be dealt with briefly later on this section, and then only in conjectural terms. The proposal below thus makes no attempt at explaining how agents are capable of attending selectively, but it does purport to subsume the phenomenon of selective attention under the semiotic framework that undergirds my theory of symbolic understanding, once again with an eye to optimizing the internal coherence of that theory. Put succinctly, I propose that selectively attending to a feature consists in interpreting that feature as a sign of a particular kind: an *icon*. Let me first briefly introduce the concept of iconicity, before turning back again to the topic of selective attention.

An icon, according to Peirce (*EP* 2:291), "is a sign which refers to the Object it denotes merely by virtue of characters of its own which it possesses, just the same, whether any such Object exists or not". Short (2007, p. 215) clarifies this as follows: "[A]n icon's significance is grounded in its own qualities, and not in any relationship to another that requires the latter to exist. (...) Anything that signifies on the ground of its own qualities alone is an icon." What to make of this? The upshot is as simple as it is mystifying: an icon is nothing but a quality that signifies itself. For instance, the specific shade of green of a given Granny Smith apple is an icon of that shade of green. That is to say, the specific shade of green, as embodied in the apple, is a sign of itself qua possibility. As Short (2007, p. 217) puts it: "The color in itself is the object; as embodied, it is the sign."<sup>98</sup> I grant that this may sound rather obscure; my clarifications further on in this section should help the reader to understand it better. All that is needed at this point, however, is to appreciate that particular items are perceived by us as having qualities (*this* apple's specific shade of green, *that* rose's sweet smell) and that Peirce regarded such qualities as signs of themselves.

Back now to our main subject: selective attention. Recall that in attending selectively, a feature comes to stand out to someone at the expense of other features. One might say that that feature is thus granted the privilege of our awareness. Awareness, however, does not entail recognition. Even though in practice the former usually gives way to the latter, the two can nonetheless be prescinded from one another. But if not recognition, what *does* selective attention amount to? What is it that our awareness of a feature consists in? The theoretical underpinnings of Peirce's account of iconicity may provide an answer to this. An icon, as said, signifies a *quality* that it is. Qualities, in Peirce's sense, are to be distinguished from other

---

<sup>98</sup> Given that signification requires (possible) responding, and because responding can only be done to something actual, icons invariably inhere in something particular. In other words, every icon needs an icon-bearer. Because of this, one might as well say that the apple itself, through embodying a particular quality, is an icon of that quality. Peirce occasionally used the term 'hypoicon' to refer to icon-bearing particulars (*EP* 2:273; Short, 2007, p. 216).

### 3 Symbolic understanding

abstracta such as properties and types. A property or type is an abstractum that comprises a continuum of possible variations: it is something general. Qualities, although abstract, are not general in this sense: they are “of precise hues or tones” (Short 2007, p. 79).

C.I. Lewis (1929, p. 121), who was intimately familiar with Peirce’s work, used the nowadays more familiar term ‘qualia’ (singular: quale) to refer to Peircean qualities. He had the following to say about them:

There *are* recognizable qualitative characters of the given (...); I call these ‘qualia’. But although such qualia are universals (...) they must be distinguished from the properties of objects. Confusion of these two is characteristic of many historical conceptions, as well as of current essence-theories. The quale is directly intuited, given, and is not the subject of any possible error because it is purely subjective. The property of an object is objective; the ascription of it is a judgment which may be mistaken; and what the predication of it asserts is something which transcends what could be given in any single experience.

A quality is abstract not because it is general, but because it lacks what Duns Scotus calls *haecceity*: “[H]ereness and nowness” (idem, p. 50). For while qualities can only ever be encountered in the here and now – namely through being embodied in particulars – their being embodied is not essential to their nature. In this sense, qualities are similar to properties and types. But unlike properties and types, qualities do not admit of variations. For this reason, there is no such thing as recognizing a quality. Recognition, after all, implies recognition *as*, which in turn implies generality in the sense of variability. To recognize something is to recognize it as a specimen of something that may be instantiated in various ways. A quality, however, being ‘one of a kind’, can only be *experienced*.<sup>99</sup> Such an experience, according to Peirce, has a quality which simply *is* the quality that is experienced. As Short (2007, p. 86) puts it: “[T]he quality we feel is the quality of our feeling.” This peculiar identity, I would suggest, is what constitutes ‘being aware of’ in selective attention: when we selectively attend, we are aware of a feature by virtue of fact that the quality of that experience is the feature attended to.<sup>100</sup>

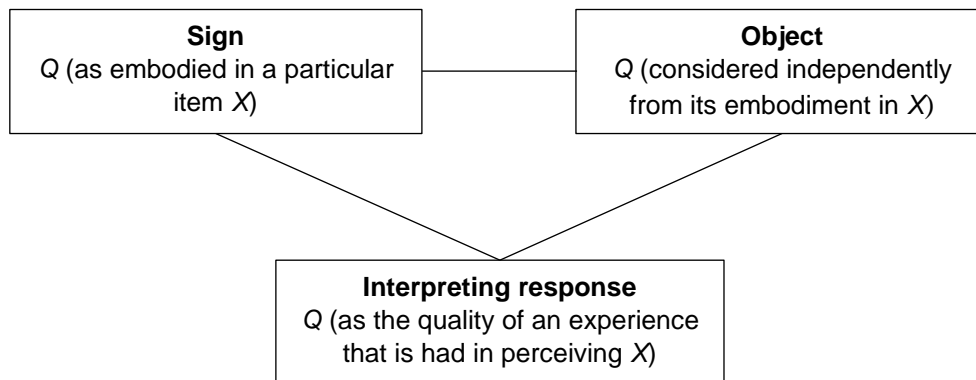
Selective attention, being awareness of a feature, can thus be analyzed in terms of quality-experience, where the notion of a quality is understood in a Peircean sense. I proposed earlier, however, that selective attention consists in interpreting a feature as an icon. In other words, the claim was made that selective attention amounts to

---

<sup>99</sup> This is not to exclude the possibility that the same quality can be experienced repeatedly; recurrence does not equate instantiation. See also footnote 64 about the difference between occurrences and instances of types.

<sup>100</sup> Note that the identity referred to is between a quality experienced and the quality of that experience: not between a quality and an experience. There can be no relation of identity between a quality and an experience, since qualities are abstracta and experiences are particular events.

subjective iconic signification. Some further remarks are required to elucidate the connection between experiencing a quality and subjective iconic signification. As explained in chapter 2, a *subjective sign* is an item that signifies something relative to an actual interpreting response. In other words, an item is a subjective sign when someone interprets that item as a sign of something else. A *subjective icon* is a type of subjective sign: namely a quality embodied in an item that signifies the quality it is experienced as being, relative to the experiential quality it is. The signifiatory triad corresponding to iconic signification looks as follows, where  $Q$  stands for a quality:



Let me try to illustrate this by means of the Granny Smith apple. Suppose you look at the apple and come to have an experience of the apple's specific shade of green. This experience has a certain quality. It is this experiential quality which constitutes a response (or more precisely: an aspect of a response) which 'interprets' that very same quality, as embodied by the apple, as a sign.<sup>101</sup> As a sign of what? Of the abstractum, the mere potentiality, that the specific shade of green is independently from its embodiment in the apple.

Recall from chapter 2 that in order for someone to interpret  $X$  as a sign of  $O$ , he must exhibit a response  $R$  to  $X$  which is subjectively justified with respect to the purpose  $R$  is produced for. With respect to what purpose(s) do experiential qualities qualify as interpreting responses that are subjectively justified? The answer is: with respect to *any* purpose. Experiential qualities always engender subjective icons, because there is no purpose on account of which they could fail to be justified.<sup>102</sup> This licenses the following inferences. Assuming that every experience of a quality has an experiential quality, and seeing that any experiential quality implicates a subjective icon, every experience of a quality breeds a subjective icon. And since the occurrence of an experiential quality entails the occurrence of an experience, every

<sup>101</sup> I use single quotation marks around 'interpret', to indicate that what we usually mean by that word stands at considerable remove from the bare 'experience-event' involved in iconic signification.

<sup>102</sup> Cf. the words of C.I. Lewis, as cited on the previous page: "The quale is (...) *not the subject of any possible error* because it is purely subjective."

### 3 Symbolic understanding

subjective icon necessitates the experience of a quality. Subjective iconic signification and quality-experience thus imply one another. Therefore, insofar as selective attention is explicable in terms of quality-experience, selective attention is explicable in terms of subjective iconic signification.

We now have a new perspective on what selective attention is: it is to ‘interpret’ a quality of an item as an icon, where a quality is a non-general abstractum embodied by an item, the nature of which consists in the nature of the experience it gives rise to. But we do not yet know what it means for a feature to be relevant to a replica’s symbolic function, nor have we gained any insight into how agents seemingly manage to have control over which features to attend to. Let me address the first issue first. When is a feature relevant to the symbolic function of a replicative sign? A replicative sign, we know from section 2.5, is an item which signifies by virtue of a rule of interpretation. And a rule of interpretation, we know too, is a conditional imperative of the form ‘If a token of *X*’s type is encountered, interpret it as a sign of *O*, or of things of *O*’s type’. Hence, something is a replicative sign if and only if it belongs to the type cited in the antecedent of some rule of interpretation. As was further pointed out at the end of section 2.5, membership of that type is not solely determined by a replicative sign’s physical properties. That is, whether the mark ‘*A*’ is an instance of the first letter of the alphabet depends not just on what it looks like, but also on the intention with which it was produced. The characteristics relevant to a replicative sign’s symbolic function are not simply certain of its physical attributes, but instead certain of its physical attributes insofar as these are exploited for a signifiatory purpose. Thus, a mark ‘*A*’ is an *A* not just if it looks like one, but rather in case the mark-producer intends her creation to be interpreted in a certain way on the basis of certain of *A*’s physical properties (e.g. stroke configuration and length). Hence, it is characteristics-cum-intentions, rather than characteristics simpliciter, which determine replicahood.

Different replicative signs may vary considerably in appearance and origin, yet possess identical symbolic functions. The characteristics-cum-intentions that determine replicahood are therefore properties, not qualities. This is because properties, unlike qualities, are general in the sense of admitting of variations. Yet in the foregoing, I talked of *features* as being relevant (or irrelevant) to symbolic function, where features were understood as qualities rather than as properties. How to repair this apparent inconsistency? Assuming we stay within Peirce’s metaphysical realm, the solution is straightforward: according to Peirce (Short, 2007, p. 79), properties are composed of qualities. Or, stated a bit more precisely, a property is a *continuum* of qualities. For Peirce, the collection of all possible – indeed, innumerable – qualities falling within a certain loosely delimited spectrum is what a property is.<sup>103</sup> This experience of greenness may be qualitatively different

---

<sup>103</sup> One should not confuse qualities with *instances* of properties. An instance of a property is something actual: it is a particular which realizes the possibility that a property is. But a

from *that* experience of greenness, but the two are akin with respect to the property they constitute: they are both qualities of which the property of greenness is composed. Therefore, a given quality, even though not by itself a determinant of replicahood, may still be said to be relevant or irrelevant to symbolic function. It is relevant in case it is constitutive of a property which determines replicahood; it is irrelevant in case it is constitutive of a property unrelated to replicahood. Among the countless qualities relevant to the symbolic function of this sentence's first letter will be qualities that are constitutive of a property relating to stroke-configuration. And among the qualities that are not relevant in this sense, will be qualities constitutive of a property relating to the depth of the letter's imprint, for instance.

So far, I have discussed only *what* selective attention might be, carefully evading the perhaps more difficult issue as to *how* agents are able to attend selectively. In the closing part of this section, I shall briefly address this issue. As there is a danger of treading too far into what is really scientific territory, my wiggle room is somewhat limited. My approach will be conjectural: based on the Peircean account of selective attention just provided, I offer a tentative proposal concerning how agents are able to attend selectively in the appropriate ways.

Selective attention is clearly an extremely fundamental phenomenon. Indeed, when explicated in terms of quality-experience as done in this section, it is the cornerstone of perception. But, although fundamental, it is not easily identified and isolated in practice. Phenomenologically speaking, quality-experiences tend to be inseparably connected to the more complex cognitive responses which they render possible. Observation, especially observation of the visual kind, almost invariably comes with implicit acts of categorization. We recognize a familiar face or an everyday object without any perceptible effort, and navigate through our surroundings by constantly, but nigh unnoticeably, ordering the environment. Difficult as it already is to monitor these subtle cognitive processes, it is even more difficult to discern the quality-experiences that underlie those processes. *A fortiori*, direct control over what qualities we experience is even harder to come by. This is also understandable in light of the fact that there is no external check on experience in and of itself: given that any purpose justifies any experience whatsoever, there is no possibility of 'friction' which could generate an impetus for exercising control. It would thus seem that selective attention is utterly undirected. However, we do have an *indirect* means for asserting control over our experiences: error at the level of categorization.

Error is our trusted companion in spurring the adjustment and refinement of largely unconscious cognitive processes. When an internal act of categorization leads to behavior that fails to meet its intended purpose, the effects of such failure

---

quality, as we saw earlier, is not something actual: although perfectly determinate, it is still an abstractum. Qualities must be embodied in order to be experienced, but their embodiment does not impinge upon their ontological status. A quality, together with uncountable quality-siblings, *constitutes* a property, an instance only *manifests* it.



### 3 Symbolic understanding

generally prompt us to change our behavior, which in turn requires changing the initial categorization. A distant relative unduly classified as a waiter at a wedding, will likely receive its proper label at a future occasion. In so adapting, selection occurs: we reduce possible variation in order to retain stability with regard to the outcomes of our doings. And such selection in turn effects selection at the level of our basal experience. For in readjusting the internal categorizations that enable certain behavior, we must readjust our experiences as well: in order to prevent ourselves from offending our poor family member again, we had better have different experiential qualities at the next encounter. While such selection at the experiential level happens more or less of its own accord, the reason why it happens in the first place is because it is implicated in a purposive process of behavioral adaptation. So, phrased in terms of the terminology from section 3.1, there is agential selection of behavior, which gives rise to a kind of mechanical selection of quality-experiences. The implication of this is that we crucially depend on our environment for taming an otherwise directionless experiential chaos. For it is by virtue of the fact that our environment provides us with the possibility of error that our behaviors are adapted and the qualities of our experiences, thereby, filtered. In case of symbol-usage, our inherent need for error is primarily catered to by other people: they impose the sanctions that help to shape our learning processes, and thus our experiences, in the ways needed for our symbol-usage to qualify as competent.

#### **3.6 Degrees of symbolic understanding: breadth**

In section 3.4, I formulated a definition of competent symbol-usage. In order to spare the reader the trouble of having to browse back through the pages, let me restate that definition here:

##### *Competent symbol-usage*

An agent  $A$  is a competent user of a symbol  $\Phi$  if and only if either

- (i)  $A$  has a disposition for using  $\Phi$  properly, and
- (ii) (i) is due to  $A$ 's having engaged in a learning process pertaining to  $\Phi$ , which was made possible by  $A$ 's having selectively attended to symbolically relevant features of one or more replicas of  $\Phi$ ,

or

$A$  would have satisfied conditions (i) and (ii) had she not been limited in her behavioral possibilities by circumstances outside of her control.

This definition, in conjunction with the semiotic framework laid out in chapter 2, already constitutes the lion's share of the explicatum of symbolic understanding. Nonetheless, there is work left to do. In particular, one important resolution made in chapter 1 has not yet been acted upon. For as I explained in section 1.4, in order to account for the scalar character of symbolic understanding we had best center the explication around comparative assessments of symbolic understanding. As of yet,

the comparative aspect does not feature in the explicatum. Over the course of two sections, I will attempt to repair this deficit, by characterizing two axes along which symbolic understanding can vary in strength.

As Kelp (2015, 3812 n13) notes, thus far few attempts have been made at developing a detailed and comprehensive account of degrees of understanding. However, there appears to be some consensus regarding the question along which *dimensions* understanding can vary in strength. In their oft-cited survey of the understanding literature, Baumberger, Beisbart & Brun (2017, 26) note that “understanding can vary in breadth and depth”, thereby presumably echoing earlier statements to that effect made by, again, Kelp (*ibidem*) as well as Elgin (2009, 324). Although claims such as these are usually made in the context of theorizing about varieties of understanding other than symbolic understanding, I will nonetheless take this consensus opinion as a starting point for formulating my own account of degrees of symbolic understanding. In doing so, I shall tailor the concepts of breadth and depth to the behaviorist and dispositionalist commitments inherent in the definition of competent symbol-usage, whilst attempting to keep my explication of the two notions sufficiently close to their common-sensical meanings. In this section, I will only touch upon the notion of breadth of symbolic understanding, leaving depth to be dealt with in section 3.7. I should like to emphasize that, in line with the nascent state of the literature on degrees of understanding, the proposal presented in this section is decidedly open-ended. Rather than providing a definitive account, I intend to offer a general framework for breadth of symbolic understanding, which allows for further details to be articulated in various ways.

On a belief-based conception of understanding, it is natural to think of breadth in terms of the comprehensiveness of a subject’s web of beliefs concerning the object of understanding: the more accurate beliefs one has, the better. How does this translate to the behavior-based account of symbolic understanding currently being developed? What makes one agent’s symbolic understanding broader than another agent’s? Here is a first stab at an answer: taking the definition of competent symbol-usage as a starting point, one could argue that comparative breadth of symbolic understanding is simply a function of how comprehensive an agent’s symbol-using competence is compared to the analogous competence of another agent. According to this idea, whether an agent may be credited with possessing broader symbolic understanding than another agent, depends on whether the former is disposed (through learning) to use a given symbol properly in a larger number of possible situations than the latter. Or, employing the prevailing terminology of the philosophical literature on dispositions, the broad understander’s disposition for proper symbol-usage has a larger set of *stimulus conditions* than the corresponding disposition of the narrow understander.<sup>104</sup>

---

<sup>104</sup> Peijnenburg (2000) offers a somewhat similar analysis of the notion of a *gradualized disposition*. For instance, the extent to which someone suffers from agoraphobia can,

### 3 Symbolic understanding

To present this first attempt at answering this section's main question in a somewhat more precise fashion, let me start by defining a basic term:

- Let  $Su_{X,\Phi}$  be the set of stimulus conditions associated with an agent  $X$ 's disposition for using a symbol  $\Phi$  properly, where that disposition was acquired through learning.

Using this term, as well as vertical bars ('|') for indicating the cardinality (=number of elements) of a set, we can define what I shall call *Numerical Breadth* as follows:

#### *Numerical Breadth*

An agent  $A$  has a broader understanding of a symbol  $\Phi$  than another agent  $B$  if and only if  $|Su_{A,\Phi}| > |Su_{B,\Phi}|$ .

A brief note on the above definition. Stimulus conditions, considered as types of situations or events, can in principle be described at any level of specificity. Consequently, in order to render the symbol-using competences of two agents commensurable, their dispositions must be assessed on the basis of the same standards for individuating events and situations. Those standards must be such that sets of stimulus conditions are finite. This important background condition will continue to operate in all of the subsequent definitions I present in this section.<sup>105</sup>

Does Numerical Breadth constitute a plausible account of breadth of symbolic understanding? To answer this question, we should assess whether its verdicts are consistent with the overall function of ascriptions of symbolic understanding.

---

according to Peijnenburg (idem, p. 296) be couched in terms of how many different types of situations or environments invoke the types of responses associated with agoraphobia.

<sup>105</sup> By taking sets of stimulus conditions to be finite, I have assumed that any way of discriminating between stimulus conditions must ultimately bottom out at some irreducible level of specificity. This is what allows comparative assessments of symbolic understanding to be couched in simple, numerical terms: namely in the sense that one's understanding of some symbol is broader to the extent that one is disposed to use that symbol in a larger number of possible situations. However, one could complain that the assumption of finitude makes for an unrealistic picture of breadth of symbolic understanding. According to this complaint, there is no upper limit to the specification of stimulus conditions: we should rather think of them as lying on a continuum. But if we take this idea on board, then simply comparing the number of stimulus conditions within two sets  $Su_{X,\Phi}$  and  $Su_{Y,\Phi}$  is no longer feasible, as both now contain an uncountable infinitude of elements. To be able to compare such sets in terms of their relative size, we would need to impose a *measure* on them. Doing this requires taking the set ( $Su_\Phi$ ) of all possible stimulus conditions pertaining to the proper usage of a symbol  $\Phi$ , and then defining a set function over its power set  $P(Su_\Phi)$  – a set function that assigns a positive real number to each and every element of  $P(Su_\Phi)$ . By stipulation, these numbers then serve to indicate the sizes of those elements. Although I am not unsympathetic to the complaint, I have chosen to stick with the assumption that sets of stimulus conditions are finite for reasons of (mathematical) convenience.

Trivially, more comprehensive dispositions for symbol-usage contribute more to upholding a stable practice of symbol-usage than less comprehensive dispositions. So a community should want to promote comprehensiveness, and Numerical Breadth seems to constitute a means for this. Further scrutiny, however, reveals that differences in the mere number of stimulus conditions for dispositions form only a crude heuristic for measuring breadth of symbolic understanding. This is because sets of stimulus conditions may be more or less homogeneous, and depending on this, less or more significant in view of the purpose of ascriptions of symbolic understanding. To argue this point, let me start with a highly simplified example. First, consider an agent whose disposition to properly use the predicate ‘is ironic’ spans exactly five types of situations. Suppose that these situations are highly dissimilar to one another: all somehow exemplify irony, but they do so in very different ways. Second, consider an agent whose disposition to properly use the predicate ‘is ironic’ also spans five situations. In this case, however, the situations are alike in the sense that they exemplify irony in similar ways. As I see it, the former agent’s disposition carries more epistemic weight than the latter’s, given that it is more robust with regard to contextual variation.

Why would robustness with regard to contextual variation be of importance to comparative assessments of symbolic understanding? My answer to this is that if ascriptions of symbolic understanding serve to promote efficient communication through maintaining stable practices of symbol-usage, then it is expedient for a community to differentially encourage the possession of heterogeneous dispositions for proper symbol-usage over the possession of more homogenous, but otherwise equally comprehensive symbol-using dispositions. This is because those with relatively heterogeneous dispositions will be able to extend their symbol-using competence with comparatively less effort. Someone who is already disposed to apply a symbol properly in a diverse set of circumstances is in a better position to extrapolate from this disposition, compared to someone who has a disposition for proper symbol-usage that comprises equally many, but unequally diverse types of situations. In an ever-changing world, in which new types of situations arise constantly, the need for agents to continuously broaden their symbol-using dispositions via extrapolation is of the essence to ensure efficient communication in the long run.

The fact that Numerical Breadth cannot accommodate the importance of heterogeneity means that we should define breadth of symbolic understanding in some other way. To do this, I start out with making the following assumption:

- Let the elements in any  $S_{u,x,\phi}$  be considered as items in a space, and let the degree of similarity between any two elements in that space be expressed in terms of their relative distance.

### 3 Symbolic understanding

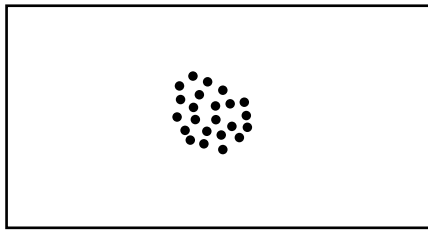
The addition of this similarity parameter provides the first step to incorporating the idea that the degree of heterogeneity of a set of stimulus conditions matters to an agent's breadth of symbolic understanding. In principle, there is more than one way of specifying how, exactly, heterogeneity contributes to breadth. The most straightforward option – and the one that I will go by – is to let an agent's breadth of symbolic understanding be determined by the volume or size of the so-called *convex hull* of  $Su_{X,\phi}$ , being the smallest enclosure of  $Su_{X,\phi}$  that contains the line segments connecting each pair of items in that set. Let us denote this volume as  $\text{conv}(Su_{X,\phi})$ . The resulting definition for what I call *Spatial Breadth* looks as follows:

#### Spatial Breadth

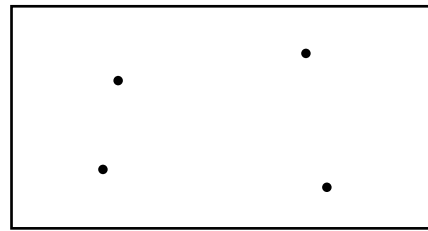
An agent  $A$  has a broader understanding of a symbol  $\Phi$  than another agent  $B$  if and only if  $\text{conv}(Su_{A,\phi}) > \text{conv}(Su_{B,\phi})$ .

On this conception of breadth, assessments of breadth of symbolic understanding become highly sensitive to inter-agent differences in the heterogeneity of symbol-using dispositions. The question is, though, whether those assessments might not become *too* sensitive to heterogeneity. For while it is important for a community to foster versatility in the proper usage of symbols, it should at the same time want to encourage *consistency*.

This requires some elaboration. Compare the following (highly simplified) spatial representations of some agent  $A$ 's  $Su_{A,\phi}$  and another agent  $B$ 's  $Su_{B,\phi}$ , in which every black dot represents a stimulus condition, and in which the distances between the dots serve to indicate how similar the stimulus conditions are to one another – relative, that is, to a common standard for measuring similarity:



Spatial representation of  $Su_{A,\phi}$



Spatial representation of  $Su_{B,\phi}$

Based on Spatial Breadth,  $B$  has a broader understanding of  $\Phi$  than  $A$ , given the fact that  $\text{conv}(Su_{B,\phi}) > \text{conv}(Su_{A,\phi})$ . It is questionable, however, whether  $B$ 's (expected) contribution to the goal of sustaining communicative efficiency is indeed greater than  $A$ 's – which should be the case if comparative ascriptions of breadth of symbolic understanding are to be tailored to that larger purpose. After all, it seems that the fact that  $|Su_{B,\phi}| \ll |Su_{A,\phi}|$  offsets the benefits of  $Su_{B,\phi}$ 's higher degree of heterogeneity. That is, the fact that  $Su_{B,\phi}$  contains so few elements compared to  $Su_{A,\phi}$  makes it less plausible to claim that  $B$ 's greater versatility puts him in a better position than  $A$  to

extend his symbol-using disposition to new cases. There is simply insufficient evidence to suppose that  $B$  possesses the consistency needed for him to successfully extrapolate from his current disposition. Hence, if a community were to systematically favor  $B$ -type dispositions over  $A$ -type ones, it would engage in a relatively risky practice of understanding-ascription. Risky, that is, with respect to the goal of ensuring communicative efficiency.

The preceding considerations go to show that while heterogeneity is important, Spatial Breadth loses out on what is plausible about the criterion encapsulated in Numerical Breadth. What we need then, is some way of combining  $|Su_{X,\phi}|$  and  $\text{conv}(Su_{X,\phi})$  into a unified measure of breadth of symbolic understanding which is sensitive to both comprehensiveness and heterogeneity. Exactly how these two parameters should be combined is far from self-evident, however. Let me give just one example to indicate why the matter is complicated. One might think, for instance, that breadth of symbolic understanding ought to be measured in terms of the product of  $\text{conv}(Su_{X,\phi})$  and  $|Su_{X,\phi}|$ . But this is too simple. For the degree to which an increase in  $|Su_{X,\phi}|$  raises one's breadth of symbolic understanding is likely to be subject to a form of the law of diminishing marginal utility. That is to say, in case  $\text{conv}(Su_{X,\phi})$  is already densely packed with stimulus conditions, the addition of one extra element to  $|Su_{X,\phi}|$  may be deemed less epistemically significant than in case  $\text{conv}(Su_{X,\phi})$  is relatively empty. But how much less significant? It is not clear what the answer should be. To be sure, this example is merely illustrative. In order to avoid entering into excessive speculation, I will not attempt to list all other possible complicating factors, nor (*a fortiori*) will I try to offer a fully specified measure of breadth of symbolic understanding that does justice to those factors.

The main takeaway from the preceding considerations, then, is that breadth of symbolic understanding hinges on two main parameters, comprehensiveness and heterogeneity, which are linked in some yet-to-be-specified manner. For ease of reference, I will dub this latest improvement over the previous definitions *Two-Dimensional Breadth*, and use the notation  $f(|Su_{X,\phi}|, \text{conv}(Su_{X,\phi}))$  to denote a function, strictly increasing in both arguments, that gives us the 'success-value' for any combination of  $|Su_{X,\phi}|$  and  $\text{conv}(Su_{X,\phi})$ :

*Two-Dimensional Breadth*

An agent  $A$  has a broader understanding of a symbol  $\Phi$  than another agent  $B$  if and only if  $f(|Su_{A,\phi}|, \text{conv}(Su_{A,\phi})) > f(|Su_{B,\phi}|, \text{conv}(Su_{B,\phi}))$

Now, independently of how  $f(|Su_{X,\phi}|, \text{conv}(Su_{X,\phi}))$  is specified, we can ask whether Two-Dimensional Breadth constitutes a complete measure of breadth of symbolic understanding. The answer is no. So far, I have only taken into consideration dispositions for *proper* symbol-usage, thus implying that breadth of symbolic understanding is solely a matter of comparative success. But apart from success, we also need to take into account failure. An agent could be disposed to use a symbol

### 3 Symbolic understanding

properly in one portion of the symbol's domain of application, but disposed to use it improperly (or not at all) in another portion. The latter fact, as well as the former, should bear on how we assess the agent's breadth of symbolic understanding.

The reason why this is so is that communities would be seriously hampered in maintaining communicative efficiency via ascribing symbolic understanding if such ascriptions did not help to eliminate failure. To only reward success, and not also curtail failure, would be to encourage the use of insufficiently risk-averse methods for acquiring dispositions for proper symbol-usage. This is because forming a disposition for proper symbol-usage may sometimes imply forming an additional disposition for improper usage of that same symbol. For instance, acquiring a disposition to properly use the predicate 'is ironic' in a certain set of paradigmatically ironic situations might, depending on the learning process underlying this disposition, automatically give rise to a further disposition for improper usage of that same predicate in a different set of superficially similar, but non-ironic situations. Were the latter disposition not discounted in an assessment of symbolic understanding, the former disposition, and thereby the learning process underlying it, would be disproportionately rewarded relative to a *slightly* weaker success-disposition which generated a *significantly* smaller failure-disposition, or even no failure-disposition at all.

Before I make an attempt at merging the evaluative parameter just introduced into the latest definition of breadth of symbolic understanding, I should clarify what I mean by failure in symbol-usage. First of all, as with success, I conceive of failure behavioristically and dispositionally: it consists in a tendency or inclination to do (or not do) certain things in certain possible situations. At the most general level, failure comes in two forms: it can either manifest itself in a *disposition for error* or in a *disposition for omission*. Dispositional error in symbol-usage amounts to being disposed to use a symbol in accordance with a purpose which is not germane to the symbol. Dispositional omission consists in being disposed to not using a symbol where proper usage would be warranted. Crucially, also, in order to count as genuine failure, dispositional error and omission in symbol-usage must result from inadequate or insufficient learning. That is, a disposition for error or omission only licenses the imputation of failure when that disposition was acquired through engaging in a faulty learning process, or by not having learnt at all.

This final proviso is important, since neither dispositional omission nor error need have repercussions for one's level of symbolic understanding per se. Dispositional omission is epistemically benign in case an agent's lack of proper symbol-usage is not traceable to some kind of learning deficit. Take, again, the use of the predicate 'is ironic'. Some agents may have a tendency to continuously (and correctly) inform others of their encounters with ironic expressions or situations to the point of annoyance; thus showcasing a wide-ranging disposition for proper usage of 'is ironic'. Others may only be disposed to use that phrase in accordance with its signifiatory purpose when, for example, explicitly asked to do so. The fact that the

latter type of agent has a disposition for omission relatively to the former type, is often epistemically inconsequential. Taciturn agents may have limited dispositions for proper usage of the word ‘irony’ (or of any word, for that matter) simply because they are temperamentally disinclined to share their otherwise accurate, private observations. And inattentive agents might have such limited dispositions merely because they are insufficiently motivated to discern irony in language or in life – perhaps they are, by disposition, constantly in a hurry and need to focus on their non-ironic business. In such cases, the taciturn’s or inattentive’s relatively restricted sets of stimulus conditions for proper symbol-usage do not implicate lower levels of symbolic understanding compared to their more talkative or observant competitors. This is because communicative efficiency as such is not jeopardized by motivationally or temperamentally restricted proper symbol-usage. A practice for symbol-usage can remain perfectly stable, regardless of whether its participants are inclined towards reticence or loquacity. Dispositional omission only constitutes failure in the epistemically relevant sense when it results from learning deficits, for such deficits *do* threaten to undermine stability.<sup>106</sup>

As said, dispositional error can likewise be devoid of epistemic import. When, for instance, an agent has certain goals which interfere with his purpose to use a symbol properly, the epistemic dimension of his performance is put between brackets, as it were. Think, for instance, of an example used in section 3.2, concerning the use of a traffic sign for covering up a hole in the floor. Evidently, this kind of usage does not constitute proper usage of the traffic sign. But neither does it necessarily amount to failure: it could well be that the hole-coverer is perfectly aware of the traffic sign’s ordinary function, but elects to use it for purposes which for some reason are of paramount importance to him. If so, there is no implication that the hole-coverer is liable to epistemically pertinent dispositional error. In a similar vein, pop musicians struggling to make the lyrics fit the rhythm of their songs, or poets trying to squeeze their words into the desired metre, sometimes tend to take some liberty in stretching the rules of grammar or the meanings of words and phrases, so as to procure an end product that better satisfies their aesthetic requirements. If an otherwise healthy symbol-using habit is overridden by such a counteracting disposition, the agent’s comparative breadth of symbolic understanding remains unaffected.<sup>107</sup>

---

<sup>106</sup> Admittedly, having a more wide-ranging disposition for proper symbol-usage makes it likelier that evidence of that disposition will be forthcoming, which in turn can raise the probability that one is credited with comparatively broader symbolic understanding by one’s fellow symbol-users. But neither availability of evidence for symbolic understanding, nor frequency of understanding-ascription, directly impinge upon an agent’s actual comparative breadth of symbolic understanding. For what matters to symbolic understanding, as was established already in section 1.5, is the *possible* rather than the *actual*: understanding resides in dispositions, not in manifestations of dispositions.

<sup>107</sup> An interesting boundary case, which also happens to attach nicely to our running example, is the song *Ironic* by Canadian singer-songwriter Alanis Morissette. As has been noted



### 3 Symbolic understanding

The rationale behind this is that agents who purposefully use symbols improperly are, as it were, momentarily stepping out of the symbol's associated practice. Communities should allow for a kind of derogation for such maneuvers for two reasons. In the first place, a community must take care not to let its need to secure stability in symbol-usage interfere unnecessarily with its other needs, such as the need to cover up dangerous holes. Secondly, a community should allow for the possibility of gradual evolution of its symbol-using practices. This latitude is needed because symbol-using practices are ultimately subservient to the worldly demands imposed upon a community – demands which are bound to change over time. For instance, sustained and intensive contact between communities without a common language has tended to necessitate linguistic creolization – a process of grammatical simplification and lexical mixing tailored to a newly arisen need to create a mutually accessible communicative practice. Such processes can only come off the ground provided that there exists some a priori leniency within communities towards purposefully improper symbol-usage. So while communicative efficiency benefits from, and is indeed predicated on, a kind of conservatism, *unrelenting* conservatism is ultimately antithetical to the proper functioning of communities. This is why, in ascriptions of symbolic understanding, purposeful improper symbol-usage is neither generally approved, nor disapproved of. Not approved of, because it does little to bolster the stability of an existing symbol-using practice. But not strictly disapproved of either, since such impropriety is, in the end, the only means a community has for keeping up with a dynamic world.<sup>108</sup>

---

widely, the lyrics to the song make reference to eleven purportedly ironic situations, none of which are, in fact, straightforward instances of irony. Whether Morissette was intentionally trying to cause confusion or stir up intellectual discussion, or whether she genuinely misunderstood the concept, remains a subject of debate. See also Roberts (2014), who argues that the examples featuring in the song's lyrics are bona fide cases of irony after all.

<sup>108</sup> To allow for gradual evolution in a community's symbol-using practices, one might think that comparative assessments of symbolic understanding should be relativized to rules of interpretation. According to this proposal, such assessments might take the following form: agent *A* is a more competent user of  $\Phi$  than *B* relative to interpretative rule  $I_1$ , but *B* is a more competent user of  $\Phi$  than *A* relative to interpretative rule  $I_2$ . In this way, one could argue, the practice of ascribing symbolic understanding would promote diversity, and thereby make more room for gradual evolution in the use of symbols. I would be opposed to such a proposal, however. For relativizing assessments of symbolic understanding to interpretative rules would threaten to defeat the main purpose of such assessments, which is to maintain communicative efficiency. Communicative efficiency necessitates (a high degree of) uniformity, which in turn necessitates compliance with a limited number of 'privileged' interpretative rules. When it comes to symbol-usage, diversity must to some extent be accommodated to allow for gradual change, but it must also be curtailed in order to prevent agents from ending up talking at cross-purposes all too frequently. In a fully relativized practice of ascribing symbolic understanding, the balance would sway too far in the direction of diversity. I should hasten to point out, however, that this does not imply that uniformity in symbol-usage had best be enforced via codified rules in a purely top-down manner. For that kind of policy tends to lead to the excesses that we know, for instance, from the writings of 17<sup>th</sup>-century Dutch grammarians such as Christiaan van Heule, who attempted to homogenize

As said, a disposition for error or omission only licenses the imputation of failure when that disposition was acquired through engaging in a faulty learning process, or by not having learnt at all. One question that might be raised in connection to this is the following: does it matter for the epistemic significance of an agent's disposition for failure whether he is in any way blameworthy for his learning deficiencies? I would answer in the negative. This means, crude as it may sound, that being disposed to make mistakes by virtue of inadequate or insufficient learning inevitably detracts from an agent's breadth of symbolic understanding, even if the agent could not reasonably have been expected to avoid acquiring that disposition. Conversely, being disposed to use symbols properly through having engaged in learning invariably contributes to one's breadth of symbolic understanding, even if the agent just happened to be in circumstances which provided him with all the required opportunities for learning. The reason why I think considerations of blameworthiness should not enter into assessments of symbolic understanding is that the aim of ensuring communicative efficiency does not leave much room for consistently exonerating non-blameworthy learning deficits. Someone whose symbol-usage is hampered by learning deficits, whatever the background of those deficits, contributes less to maintaining stable practices of symbol-usage than someone whose symbol-usage is not so restricted. To be sure, a community must strive to forestall and remove any structural inequalities which may give rise to involuntary learning deficits. But it is not to this goal that the practice of ascribing symbolic understanding is geared. In the next section, I make a similar point with regard to depth of symbolic understanding.

We can now make an attempt at building the notion of failure into our definition of breadth of symbolic understanding. I propose that this can be done by conceiving of breadth as success minus failure. In order to make this somewhat more precise, let me first define another term:

- Let  $Fa_{X,\Phi}$  (with 'Fa' for 'failure') be the set of stimulus conditions associated with an agent  $X$ 's disposition for error and/or omission regarding the use of a symbol  $\Phi$ , where that disposition was acquired through inadequate or insufficient learning.

As I explained previously, heterogeneous success-dispositions are, by and large, more commendable than homogeneous ones, given that the former are more readily extrapolated from than the latter. Yet I also indicated that the positive epistemic impact of heterogeneity can be offset by decreased consistency. When we apply this kind of thinking to the notion of failure, the following picture emerges. First, it can

---

the (by then already obsolete) Dutch case system with the Latin case system. Such rigidity undermines, rather than bolsters, communicative efficiency. In the end, which interpretative rules count as privileged in a given time and context, is determined by intractable social developments – developments which standardizers can only marginally interfere with.

### 3 Symbolic understanding

be noted that heterogeneous failure-dispositions generally detract more from an agent's breadth of symbolic understanding than homogeneous ones, since the former are more prone than the latter to generalize to further types of situations. But the negative epistemic impact of heterogeneity can be neutralized by a lack of saturation within the failure-disposition. That is, if one's disposition for failure contains proportionally few stimulus conditions across a wide variety of possible situations, then that disposition is relatively benign compared to one which contains far more stimulus conditions within a slightly narrower range of possible situations.

Factoring this in, and relying again on the conception of similarity between types of situations in terms of distances between items in a space, we can now update the definition of Two-Dimensional Breadth. Mirroring earlier decisions, I remain neutral as to how exactly comprehensiveness and heterogeneity relate when it comes to failure. Indeed, I remain non-committal as to whether comprehensiveness and heterogeneity relate to one another in the same way that they relate to one another in case of success. To reflect this choice, I use  $f^*(|Fa_{X,\phi}|, \text{conv}(Fa_{X,\phi}))$  to denote a function, again strictly increasing in both arguments, that gives us the 'failure-value' for any combination of  $|Fa_{X,\phi}|$  and  $\text{conv}(Fa_{X,\phi})$ . The new definition of breadth, called *Net Two-Dimensional Breadth*, looks as follows:

#### *Net Two-Dimensional Breadth*

An agent  $A$  has a broader understanding of a symbol  $\Phi$  than another agent  $B$  if and only if

$$(f(|Su_{A,\phi}|, \text{conv}(Su_{A,\phi})) - f^*(|Fa_{A,\phi}|, \text{conv}(Fa_{A,\phi}))) > (f(|Su_{B,\phi}|, \text{conv}(Su_{B,\phi})) - f^*(|Fa_{B,\phi}|, \text{conv}(Fa_{B,\phi})))$$

Does this provide us with an accurate account of breadth of symbolic understanding, then? Alas, again, no. As it stands, Net Two-Dimensional Breadth does not fully account for the fact that  $|Su_{X,\phi}|$  and  $\text{conv}(Su_{X,\phi})$  can vary depending on factors other than ones related to learning. That is, in its current form the definition unduly disadvantages those who, through sheer motivational or temperamental factors, have less comprehensive success-dispositions than others.

This shortcoming of the definition is due to the fact that while being disposed to use a symbol properly in *fewer* possible situations does not necessarily detract from one's level of symbolic understanding (i.e. in case of non-learning-related omission or error), being disposed to use a symbol properly in *more* possible situations does in fact always add at least something to one's level of symbolic understanding. To correct for the unfortunate implications of this 'more-is-better-principle', the taciturn and the inattentive must be calibrated with the talkative and the observant. That is, in order to avoid that epistemic inequalities can result from mere motivational or temperamental differences, we have to bring an artificial equalizing instrument to the table. What instrument should this be? I suggest it can be retrieved from this very

dissertation: the notion of a *counterfactual disposition*, as used in section 3.2. A counterfactual disposition is a disposition that would have been actual, had certain circumstances been different. Using this concept, those seemingly lacking in breadth can be brought on a par with their ostensibly better performing counterparts through redirecting the locus of epistemic assessment to a possible world in which they, in a sense, *are* their counterparts. In this way, a community has a tool for securing stability in symbol-usage which does not have the undesirable collateral effect of pressuring the community's population into motivational and temperamental uniformity.

Now, of course, agents are rarely unidirectionally motivated or predisposed across the board. Some may be prone to display their symbol-using competence before a certain public, but inclined to being highly reticent elsewhere. Talkativeness, attentiveness, reticence and absent-mindedness tend to be context-dependent traits. The correction mechanism must therefore be such that for each type of situation relevant to a comparative assessment of symbolic understanding, the mechanism maximizes the breadth of symbolic understanding for both agents relative to one another. This can be achieved by defining, in addition to  $Su_{X,\phi}$  and  $Fa_{X,\phi}$ , a third set:

- Let  $CSu_{X,\phi}$  (with 'CSu' for 'corrected success') be the set of stimulus conditions associated with the maximally wide-ranging learning-based disposition for using a symbol  $\Phi$  properly, which agent  $X$  would have possessed had  $X$  been similarly motivated or otherwise non-epistemically inclined as another agent  $Y$  in those situations which, in the actual world, constitute stimulus conditions for  $Y$  but not for  $X$ .

Thus, in somewhat more precise terms,  $CSu_{X,\phi}$  is determined as follows:

- For any pair of sets  $Su_{X,\phi}$  and  $Su_{Y,\phi}$ ,  $CSu_{X,\phi}$  covers all elements in  $Su_{X,\phi}$ , as well as all elements in  $Su_{Y,\phi} \setminus Su_{X,\phi}$  which are not absent from  $Su_{X,\phi}$  for reasons having to do with insufficient or inadequate learning, and which would have been in  $Su_{X,\phi}$  had  $X$ , in the stimulus conditions at issue, been similarly non-epistemically predisposed as  $Y$ .

The considerations concerning heterogeneity and homogeneity apply as they did before. This means that the updated definition, labelled inelegantly as *Net Corrected Two-Dimensional Breadth*, looks as follows:

*Net Corrected Two-Dimensional Breadth*

An agent  $A$  has a broader understanding of a symbol  $\Phi$  than another agent  $B$  if and only if

$$(f(|CSu_{A,\phi}|, \text{conv}(CSu_{A,\phi})) - f^*(|Fa_{A,\phi}|, \text{conv}(Fa_{A,\phi}))) > \\ (f(|CSu_{B,\phi}|, \text{conv}(CSu_{B,\phi})) - f^*(|Fa_{B,\phi}|, \text{conv}(Fa_{B,\phi})))$$

It is worth pointing out that  $Fa_{X,\phi}$  does not require a similar modal extension mechanism as  $Su_{X,\phi}$ . This is because of an inherent asymmetry between the notions of success and failure (as understood here): inter-agent differences in motivation and character must be corrected for when it comes to comparative success, but not when it comes to comparative failure. After all, dispositions for error or omission brought about by motivational or characterological factors are, by virtue of how I have defined failure, already deleted from the picture.

In this section, I have theorized about one dimension of symbolic understanding's graduality: breadth. In keeping with what I said at the outset, I am not claiming that the proposal I have presented is the only possible way to conceptualize breadth. I do believe, however, that the proposal contains two key components that should figure in any account of breadth of symbolic understanding: namely the parameters of comprehensiveness and heterogeneity. Furthermore, the need to factor in failure (besides success), and to correct for epistemically irrelevant differences among agents, are aspects that I believe any theory of breadth should do justice to. Be that as it may, the endpoint of this section's inquiry, the definition of Net Corrected Two-Dimensional Breadth, contains important 'gaps' that have yet to be filled in in some way. These gaps concern, first of all, the exact way in which comprehensiveness and heterogeneity contribute to breadth, and secondly, the way in which failure detracts from success. It is thus apt that my account of breadth be viewed as considerably more open-ended than most of the conclusions arrived at elsewhere in this dissertation. Something similar goes for the next section, which deals with the second dimension of symbolic understanding's graduality: depth.

### 3.7 Degrees of symbolic understanding: depth

The first criterion of my account of symbolic understanding dictates that in order to understand a symbol, an agent must be disposed to use that symbol's replicas in accordance with their purpose. In the previous section, I proposed a way of gradualizing this criterion. According to the second criterion associated with my account, the agent must have acquired her disposition through having engaged in a learning process relevant to the symbol at hand. This criterion leaves room for gradualization too. As I hope to show in the present section, gradualizing the second criterion provides us with a measure of *depth*. As with breadth, the contemporary literature on understanding offers little clues as to how we should define depth for understanding in general, let alone how we should define it for a behaviorist and dispositionalist conception of symbolic understanding. With respect to the account of breadth developed in the previous section, I have thus had to work more or less from scratch. This mode of operation will continue below. Still, the analysis that I shall offer is not a complete shot in the dark. I will take my cue from what has been

### 3.7 Degrees of symbolic understanding: depth

said about depth of explanation by philosophers of science, applying the basic idea behind a familiar view of explanatory depth to the concept of symbolic understanding. More specifically, I shall argue that an agent's depth of symbolic understanding is a function of how cognitively economical his learning processes are. The more general aim of this section is, as before, to improve upon the explicandum of symbolic understanding in terms of its exactness.

Although detailed treatments of depth are all but lacking in the literature on understanding, some of the scattered remarks that have been made may nonetheless provide us with useful input for exploring new ground. For instance, Kelp (2015, 3812n13) has suggested that an agent's depth of understanding of a phenomenon may be associated with the 'well-connectedness' of his knowledge concerning that phenomenon. An agent's knowledge is *maximally* well-connected just in case "the basing relations that obtain between the agent's beliefs about P reflect the agent's knowledge about the explanatory and support relations that obtain between the members of the full account of P" (idem, 3810). In a somewhat similar but much more cursory vein, Elgin (2009, 324) has claimed that depth is a function of how "tightly woven", a subject's web of beliefs is. The common denominator between these two skeletal accounts of depth seems to be *density* of a sort: the denser the structure of interrelations within a subject's collection of beliefs concerning some object of understanding, the deeper his understanding of that object.

Interestingly, the density-view of depth of understanding markedly differs from what has been written about depth in the philosophy of science literature. As explained in section 1.3, philosophers of science tend to think of understanding as being either constituted or mediated by sound explanations. On this picture, depth of understanding equates, or is at least strongly correlated with, depth of explanation. In an article devoted to this topic, Weslake (2010, 283) points out that according to two of the most influential accounts of scientific explanation – Hempel's DN model and Woodward's so-called interventionist model – *generality*, rather than density, is the mark of explanatory depth, and hence of depth of understanding: "While the (...) DN view defines depth in terms of the range of possible systems to which an explanatory generalization potentially applies (scope), the interventionist view defines depth in terms of the range of counterfactual questions an explanatory generalization answers (invariance)." Explanatory depth is thus conceived of as a function of the unifying power of the explanans. In what follows, I shall argue that we can identify a second dimension of graduality in symbolic understanding which bears a clear, if only structural, resemblance to the generality view of depth.

In chapter 2, I formulated a definition of symbolicity which was meant to cover a wide range of representational devices: letters, words, pictograms, sentences, gestures, diagrams, physical models, and so on. According to this definition, any item which signifies something on the basis of a rule of interpretation is a symbol. Now, within this motley collection of items, a distinction can be drawn between

### 3 Symbolic understanding

*atomistic*, *molecular* and *composite* symbols. Here are the three definitions corresponding to these concepts:

*Atomistic symbol*

A symbol  $\Phi$  qualifies as atomistic ( $\Phi_A$ ) when its replicas are wholly composed of non-symbolic elements.

*Molecular symbol*

A symbol  $\Phi$  qualifies as molecular ( $\Phi_M$ ) when its replicas are partly or wholly composed of symbolic elements  $\varphi_1, \dots, \varphi_n$  whose symbolic functions are not part of what constitutes the symbolic function of  $\Phi$ .

*Composite symbol*

A symbol  $\Phi$  qualifies as composite ( $\Phi_C$ ) when its replicas are partly or wholly composed of symbolic elements  $\varphi_1, \dots, \varphi_n$  whose symbolic functions jointly determine the symbolic function of  $\Phi$ .

In the present context, the category of composite symbols is of special interest. For as I hope to show, it is with regard to composite symbols (and to composite symbols only) that agents can have deeper or shallower understanding. More specifically, I submit that when it comes to understanding such symbols, agents may differ in their comparative depths of symbolic understanding on account of how their dispositions for proper usage of the symbol were formed. The more *cognitively economical* the process of disposition-acquisition, the deeper the agent's understanding of the composite symbol.

Before proceeding, it is apt to adduce some examples in order to illustrate the three concepts just defined. First, instances of atomistic symbols are for example the letters of the Latin alphabet, the dashes and dots of Morse code, as well as simple types of facial expressions such as raised eyebrows. Such items constitute the smallest meaningful elements within our various communicative repertoires; further analysis into more basic constituents will reveal only items that are non-semiotic in nature (e.g. lines or pixels, muscular contractions). Molecular symbols, secondly, are similar to atomistic ones in the sense that both harbor an aspect of irreducibility: as with atomistic symbols, the symbolic function of molecular symbols is not determined by the symbolic functions of their parts. Unlike atomistic symbols, however, molecular symbols do contain symbolic elements. Examples of molecular symbols are the majority of non-compound words of natural languages, as well as most idiomatic expressions. Both are made up of symbolic elements (letters/morphemes and words/phrases, respectively), but – generally speaking – those elements do not jointly determine the symbolic function of the symbol of which they are part. Finally, composite symbols are those entities whose symbolic functions are more or less directly derivative from the symbolic functions of their

parts. Sentences with non-figurative meanings provide a case in point, but so do many diagrams and physical models. The significatory function of such symbols can be analyzed in terms of the significatory functions of their parts in conjunction with rules for combining those parts.

Further on in this section, I will come back to the threefold distinction just introduced and illustrated, and highlight a couple of nuances. For now, let me turn to explaining what I think depth of symbolic understanding consists in. I said earlier that depth of symbolic understanding pertains to the understanding of composite symbols, and that an agent's comparative depth in understanding a composite symbol is proportionate to how *cognitively economical* her process of acquiring a disposition for using the symbol was. When is a disposition acquired in a cognitively economical fashion? At the most general level, a disposition for proper usage of a composite symbol can be acquired in two ways: either through a *direct*, or through an *indirect* learning process. I speak of *direct learning* when the cognitive mechanism involved in the learning process (be it prototype-, exemplar-, or rule-based) takes as its input replicas of the symbol to which the disposition pertains. And I speak of *indirect learning* in case the cognitive mechanism takes as its input replicas of symbols which are elements of the symbol under consideration. Starting from this distinction, I propose that a learning process is cognitively economical in proportion to how closely it approximates a *maximally indirect* learning process.

To specify this idea a bit further, consider two agents who are compared in terms of their understanding of some composite symbol  $\Phi_C$  which consists of the symbolic elements  $\varphi_1$ ,  $\varphi_2$  and  $\varphi_3$ . Suppose that  $\varphi_1$ ,  $\varphi_2$  and  $\varphi_3$  jointly constitute the maximally reductive analysis of  $\Phi_C$  into simpler symbolic elements – any attempt at identifying a yet deeper analysis would no longer yield elements which can be said to co-determine the symbolic function of  $\Phi_C$ .<sup>109</sup> Given these assumptions, an agent can acquire a disposition for proper-usage of  $\Phi_C$  either on the basis of replicas of  $\Phi_C$  itself, on the basis of replicas of its simplest constitutive elements  $\varphi_1$ ,  $\varphi_2$  and  $\varphi_3$ , or on the basis of replicas of elements individuated at an intermediate level of analysis (e.g.  $\varphi_{12}$  and  $\varphi_3$ , or  $\varphi_1$  and  $\varphi_{23}$ ). The more the agent's manner of learning accords with  $\Phi_C$ 's deepest level of analysis, the more indirect her learning is.

Let me make this more concrete. Suppose that  $\Phi_C$  is the declarative sentence 'Sarah is John's mother-in-law' and assume that there are two agents, Anna and Bill, who have equally broad understandings of this sentence (see the definition of breadth in section 3.6). Now, even if this is the case, there may nevertheless be an important difference between the two. Suppose first that Bill has acquired his disposition to use the sentence properly by learning it in its entirety. This would mean that he observed

---

<sup>109</sup> What matters to an analysis being 'maximally reductive' is hence whether it identifies the most fundamental symbolic components of  $\Phi_C$  which still contribute to  $\Phi_C$ 's symbolic function. This means that the elements  $\varphi_1$ ,  $\varphi_2$  and  $\varphi_3$  may be either atomistic or molecular symbols; there is no implication that a maximally reductive analysis must always feature only atomistic symbols.



### 3 Symbolic understanding

one or more replicas of the sentence in the past – observations which consciously or unconsciously set in motion a learning mechanism such that Bill came to internally categorize the observed items as replicas of the sentence ‘Sarah is John’s mother-in-law’. Suppose further that, unlike Bill, Anna’s disposition for proper usage of the sentence sprang from learning its main constituents piecemeal. One can imagine, for instance, that Anna learnt ‘Sarah’, ‘is’, ‘John’s’ and ‘mother-in-law’ (or even ‘mother’ and ‘...-in-law’) separately, and acquired her disposition on the basis of those separately categorized elements. Bill and Anna, although equally well-versed in using the sentence, thus differ with regard to how they learnt it: Anna’s learning process was more indirect than Bill’s.

I take it that many will share the intuition that in this scenario, Bill’s understanding is in a way more superficial, or parrot-like, than Anna’s; Anna has somehow ‘carved deeper’ into the sentence than Bill. But this is not the main reason why I want to say that Anna has a higher degree of symbolic understanding than Bill. Instead, my argument for the claim that indirect learning is more epistemically commendable than direct learning, is that it allows agents to do more with less: relatively to the direct route, the indirect route enables agents to acquire a broader set of symbol-using dispositions. This is because symbolic elements are replicable in more ways compared to entire composite symbols. While replicas of composite symbols only occur by themselves or as part of replicas of larger composite symbols, elements of composite symbols are bound to occur in various other symbols as well. The sentence ‘Sarah is John’s mother-in-law’ is replicated in occurrences of that sentence and in occurrences of sentences such as ‘Sarah is John’s mother-in-law and William is John’s father-in-law’. But the proper name ‘Sarah’ is replicated in those sentences, as well as in many others (‘Hi Sarah!’). Learning the entire sentence is thus bound to provide one with a more limited symbol-using competence compared to learning it piecemeal. Communities should want to foster the kind of frugality associated with indirect learning, given that it reduces the time and effort needed for agents to learn symbols. After all, since these resources are invariably at least somewhat scarce, the fact that indirect learning makes a comparatively small claim upon those resources renders it a more effective instrument for maintaining a given degree of communicative efficiency than direct learning.

There is a structural similarity between the generality view of explanatory depth on the one hand, and the cognitive-economy view of depth of symbolic understanding on the other. Explanations which are general apply to a wider range of possible systems and answer a larger set of counterfactual questions pertaining to the explanatory target. And learning mechanisms which take as their input replicas of elements of composite symbols will effect more wide-ranging competences than learning mechanisms which take as their input replicas of entire composite symbols. But this is as far as the likeness goes. For the arguments for both views differ. Proponents of the generality view of explanatory depth tend to appeal to the intrinsic value of unifying power or simplicity. By contrast, my main argument for the

cognitive-economy view of depth of symbolic understanding is that a community should want to promote indirect learning processes over direct ones, given that the former make a proportionally larger contribution to the goal of maintaining communicative efficiency.

Now that the notion of depth of symbolic understanding has been clarified, one might ask – in light of what was done in the previous section – whether depth can be defined in more precise terms. I think this is possible, provided that we first devote some more attention to the distinction between atomistic, molecular and composite symbols. For admittedly, the demarcation of these three categories is somewhat less straightforward than suggested earlier. For instance, are compound words always composite, and non-compound ones always molecular? It depends. Non-compound words such as those generated through morphological derivation can sometimes be treated as composite symbols in case, for instance, the word's root and its derivational morpheme clearly contribute independently to the symbolic function of the word. Think for instance of the adverb 'gladly', the meaning of which recognizably derives from the semantically corresponding adjective 'glad' through addition of the suffix '-ly'. Conversely, compound words do not always qualify as composite symbols. When, for instance, the meaning of the compound is not an obvious 'sum' of the meanings of its constituent words (e.g. the word 'breakfast'), it is sometimes best subsumed under the category of molecular symbols. A further demarcation problem surfaces within the domain of idiomatic expressions. Most idiomatic expressions resist an exhaustive analysis into smaller semantic elements that jointly determine the expression's meaning. Then again, a certain idiom may be relatively etymologically transparent, thus enabling the language-user to identify the sub-idiomatic words and phrases contributing to the idiom's overall meaning.<sup>110</sup> It is to be expected that in light of these nuances, different symbol-users may categorize symbols (i.e. as composite, molecular or atomistic) in different ways.

The main takeaway from these examples is that if we want to offer a definition of depth of symbolic understanding, we must somehow account for the fact that there can be multiple defensible ways of delineating the categories of atomistic, molecular and composite symbols. I propose that this can be done by relativizing assessments of depth to *hierarchies of symbolic analyses*. A hierarchy of symbolic analyses consists in a set of possible analyses of a composite symbol  $\Phi_C$  into constitutive elements, which is ordered in terms of how comparatively reductive those analyses are. By relativizing I mean that when we compare two agents on the basis of their respective depths of understanding a symbol  $\Phi_C$ , that comparison must take place against the background of a single hierarchy of symbolic analyses. Taking this idea on board, let me now formulate a definition of depth of symbolic understanding. Here are some preliminaries:

---

<sup>110</sup> See also Nunberg, Sag & Wasow (1994) for a discussion of the compositionality of idioms.

### 3 Symbolic understanding

- Let  $an(\Phi_C)_X$  be the analysis of some  $\Phi_C$ , corresponding to an agent  $X$ 's input to her cognitive mechanism for learning  $\Phi_C$ .
- Let  $M$  be a set of possible analyses of  $\Phi_C$  into constitutive elements.
- Let  $>_{red}$  be a relation, defined as follows:  $x >_{red} y := x$  is reducible to  $y$ .
- Let  $>_{red}$  be an order on  $M$  such that  $(M, >_{red})$  is a well-ordered set, meaning that for any two analyses  $an(\Phi_C)_n$  and  $an(\Phi_C)_m$  in  $M$ , either  $an(\Phi_C)_n >_{red} an(\Phi_C)_m$  or  $an(\Phi_C)_m >_{red} an(\Phi_C)_n$ .

Using these formalisms, depth of symbolic understanding can be defined thus:

#### Depth

An agent  $A$  has a deeper understanding of a composite symbol  $\Phi_C$  than another agent  $B$  relative to  $(M, >_{red})$ , if and only if given  $an(\Phi_C)_A \in M$  and  $an(\Phi_C)_B \in M$ , it is the case that  $an(\Phi_C)_B >_{red} an(\Phi_C)_A$ .

Somewhat more informally stated, this definition says that  $A$  has a deeper understanding of  $\Phi_C$  than  $B$  just in case the analysis of  $\Phi_C$  which undergirds  $A$ 's learning of  $\Phi_C$  is reductively prior to the corresponding analysis pertaining to  $B$ , relative to some hierarchy of possible analyses of  $\Phi_C$  to which both  $A$ 's and  $B$ 's analyses belong.

I close off by briefly addressing two issues that may have bothered the reader. In the previous section, I asked whether considerations of responsibility and blameworthiness should impinge upon assessments of breadth of symbolic understanding. For instance, does it matter to an agent's breadth of understanding whether he is to blame for any learning deficiencies he may have? I answered no, saying that inadequate learning has negative epistemic repercussions regardless of whether the agent can be said to bear responsibility for his cognitive shortcomings. The ascriber of symbolic understanding is a strict taskmaster, in that respect. With regard to the topic of the present section, one could ask a similar question: if agents can reasonably be excused for their failure to engage in an indirect learning process, should this influence how we assess their depth of symbolic understanding? My verdict is no different this time: considerations of responsibility have no bearing upon ascriptions of depth. And although this conclusion may sound somewhat harsh, I stand by it – for the same reason that I resisted accommodating non-blameworthy learning deficiencies in the previous section. A community that wants to have a high level of communicative efficiency must be prepared to differentially promote the possession of those dispositions which contribute most to reaching and maintaining that high level. Consistently correcting for non-blameworthy shortcomings puts a significant break on this mechanism, in the sense that it reduces the extent to which

ascriptions of symbolic understanding track how strongly agents contribute to securing communicative efficiency.<sup>111</sup>

Secondly, the reader may have noticed that my account of depth does not say that an agent's symbolic understanding is deeper to the extent that the agent is more reflectively aware of, or indeed has more accurate beliefs about, the compositional nature of the symbol he is disposed to use properly. This omission is intentional, for two reasons. First of all, depth in understanding symbols, as I conceive of it, pertains to an agent's manner of learning, and learning, as explained in section 3.4, in no way presupposes or implicates conscious appreciation of what is learnt. And secondly, insofar as one may have an intuition that awareness of compositionality should positively affect an agent's depth of symbolic understanding, I would explain away this intuition by arguing that to reflectively appreciate how a symbol's meaning derives from its parts amounts to a different kind of understanding: *factual understanding*. In the next chapter, I demarcate this concept in negative terms: factual understanding concerns any kind of understanding which either targets a non-symbolic object, or a symbolic object which is not conceived of as such. From this demarcation criterion, I then infer that understanding facts about symbols (e.g. facts concerning a symbol's compositional nature) falls within the purview of factual understanding.

### 3.8 Conclusion

What is it to understand a sentence, gesture, diagram or model? Answering this question constituted one of the two main goals of this investigation. In chapter 1, I laid out the method to be used in pursuit of that goal: Carnapian explication. My aim was to improve upon the pretheoretic notion we have of symbolic understanding – the explicandum – by rendering that explicandum more exact, as well as fruitful for explicating factual understanding. In doing so, I sought to maintain a maximally broad range of applicability for the explicatum. As a first step to carrying out these tasks, chapter 2 offered a very general definition of symbolcity on the basis of C.S. Peirce's theory of signs. With this definition in hand, I have in this chapter carried out the remaining part of the explication: spelling out the notion of competent

---

<sup>111</sup> In principle, one could counter this by saying that ascriptions of symbolic understanding ought not to be geared towards securing optimal communicative efficiency, but towards securing equity in the amount of effort agents invest in participating in a communicative practice. In this dissertation, however, I have ventured to account for symbolic understanding's *actual* function, not for some function it *should* have. This descriptive approach, I have argued, implicates a negative answer to the question whether non-blameworthy learning deficits should be accommodated in making assessments of depth of symbolic understanding. By and large, everyday experience vindicates this conclusion. When, for instance, a teacher judges a pupil's understanding of a foreign language, he will tend to look simply at what the pupil is capable of in using that language. Sentiment-wise, the teacher's attitude towards his pupil's failures might be more or less forgiving depending on how he assesses the pupil's starting predicament and general capabilities, but this attitude will in all likelihood not affect the grade the pupil receives for his exams.

### 3 Symbolic understanding

symbol-usage. First, symbol-usage was defined as the selection of a symbol for a purpose. I then showed that the kind of symbol-usage of interest within the current context is proper symbol-usage: the selection of a symbol for the purpose for which it exists. As I subsequently sought to explain, however, propriety does not equate competence: competence also requires that one's disposition for proper symbol-usage is appropriately formed. Formed how, exactly? I proposed the following: an appropriately formed disposition for proper symbol-usage is one whose presence is due to a learning process, where that learning process was made possible by the agent's having attended selectively to symbolically relevant features of replicas. I tried to clarify this latter requirement by connecting the phenomenon of selective attention to Peirce's idea of iconicity, and by suggesting how selective attention may come about. In the final part of the chapter, I identified and defined two dimensions of graduality in competent symbol-usage: breadth and depth.

The explicatum of symbolic understanding I have constructed has served to account for the function of ascriptions of symbolic understanding. In the next chapter, I demonstrate its fruitfulness by showing that it constitutes an important component of my explicatum of factual understanding.

## Chapter 4: Factual understanding

In chapter 1, I defined the notion of objectual understanding in terms of the construction ‘to understand + noun (phrase)’. Within this grammatically delineated category, I subsequently identified two further, epistemological categories: symbolic and factual understanding. The former concept has been the focal point of the previous chapter, where I endeavored to explicate symbolic understanding in terms of the competent use of symbols, using the definition of symbolicity that was formulated in chapter 2. As explained in the opening chapter, the fruits of this labor were meant to be reaped in an indirect way: namely through showing how an account of factual understanding could be built around the explicatum of symbolic understanding. In this chapter, I will take up this task. The upshot of my proposal will be that factual understanding is constituted by dispositions for transmitting information relating to the environment agents find themselves in. I shall argue that so conceived, factual understanding hinges upon symbolic understanding, since the transmission of information must happen via the competent use of symbols. Given that factual understanding thus depends on symbolic understanding, the benefits of my account of factual understanding – which I hope to demonstrate the existence of – accrue indirectly to my explicatum of symbolic understanding.

The chapter consists of the following sections. In **section 1**, I demarcate factual understanding in more precise terms, offer a brief overview of contemporary views about factual understanding, and suggest a way of explicating that concept in a way that runs parallel with my explication of symbolic understanding. In **section 2**, I flesh out the broad-strokes proposal from section 1 and show that a definition of factual understanding must include my explicatum of symbolic understanding as a necessary condition. The ensuing three sections revolve around the defense of the proposal. **Section 3** seeks to clear away some possible sources of confusion looming in my definition of factual understanding. **Section 4** tries to counter a number of objections that could be levelled at that definition. And **section 5**, finally, lists the benefits of my account of factual understanding. Those benefits are mostly discipline-internal: they concern novel, or improved solutions to problems that epistemologists have grappled with in their attempts to develop theories of understanding.

### 4.1 Towards an explication of factual understanding

In my effort to explicate symbolic understanding, I have in the previous chapters been concerned mainly with three of the four requirements associated with Carnapian explication: simplicity, exactness, and similarity to the explicandum. Let me briefly summarize the progress that has been made on this front. First, simplicity. This criterion, understood in terms of conceptual parsimony, motivated the search for a definition of symbols which could capture a seemingly heterogeneous collection of types of representational devices (chapter 2). In addition, considerations

#### 4 Factual understanding

of conceptual parsimony prompted my efforts to develop accounts of agency (section 2.1) and use (section 3.1) that would dovetail with the Peircean semiotic framework. Second, exactness. The aim of improving upon the explicandum's (lack of) exactness constituted an important concern in spelling out the notions of usage, propriety and competence (sections 3.1-3.4). This aim also stood at the forefront in my attempt to gradualize the first two conditions of my definition of competent symbol-usage (sections 3.6 and 3.7). Third, similarity. On the one hand, a base-line measure of similarity to the explicandum was preserved through retaining the dispositionalist commitment implicit in our intuitions concerning symbolic understanding (section 1.5). On the other hand, I also made a considerable concession: I indicated that while we intuitively take symbolic understanding to harbor a behavioral as well as a cognitive dimension, my proposal would explicate symbolic understanding in exclusively behavioral terms.

The guiding idea behind the method of explication is that we can sometimes legitimately depart from a term or concept's standard meaning if doing so yields an understanding of that term or concept which better serves certain of our intellectual or practical purposes. In the present context, the question is how the 'considerable concession' mentioned above is to be compensated for. I am thus confronted with the task of making good on Carnap's fourth requirement: fruitfulness. What is needed, exactly, for this requirement to be met? Carnap himself allowed his readership some leeway in interpreting the notion of fruitfulness, issuing only general prescriptions such as the following:

The explicatum is to be a fruitful concept, that is, useful for the formulation of many universal statements (empirical laws in the case of a nonlogical concept, logical theorems in the case of a logical concept). (Carnap, 1950, p. 7)

Monetizing on this interpretative freedom, various authors have tailored the notion of fruitfulness to their respective intellectual ends, ranging from the purely philosophical to the political. A common tenet among these various interpretations, however, is that they explicitly or implicitly adhere to the principle that "an explication is useful or fruitful when it delivers 'results' that could not be delivered otherwise (or with much more difficulty), i.e. with the explicandum alone" (Dutilh Novaes & Reck, 2017, p. 206). I likewise adopt this principle. As indicated previously, the fruitfulness of my explication of symbolic understanding resides in the fact that it enables us to formulate an explicatum of factual understanding which provides answers to certain theoretical issues that have featured prominently in the epistemological debate on understanding. Per the above principle, the explicata must be necessary, or at least comparatively better suited, for obtaining the intended results.

In this section, I highlight some of what has been said about factual understanding in the contemporary epistemological literature, and subsequently outline the

#### 4.1 Towards an explication of factual understanding

approach I will take towards explicating factual understanding. This obliges me, however, to first say a bit more about the distinction between symbolic and factual understanding. Recall that I said in section 1.1 that factual understanding covers those forms of objectual understanding (i.e. ‘to understand’ + a noun phrase) which cannot be subsumed under the heading symbolic understanding. Let me now specify this negative demarcation principle in terms of the following, disjunctive criterion:

*Demarcation criterion for factual understanding*

A noun(-phrase)-denoted object is targeted by factual understanding if and only if either

- (1) The object is neither an entity with a symbolic function nor a system of such entities, or
- (2) The target is either an entity with a symbolic function or a system of such entities, but in the context at hand it is not being considered as such.

Disjunct (1) entails that someone has factual understanding when she understands a person, a historical event, a phenomenon, a natural kind, etc. Such objects of objectual understanding, although they may be symbolically represented, are not themselves symbolic and are therefore not targeted by symbolic understanding. Of course, we may understand them by means of, or in terms of, a symbolic representation, but that is a separate matter which I will elaborate on further on in the present section, as well as in the next.

Disjunct (2) says that one can have factual understanding of symbols, or systems of symbols, in contexts where those symbols or symbol systems are treated as belonging to a non-symbolic category. For instance, one can have understanding of a word through grasping its etymological origin, and one can understand a language by seeing how it relates to, and diverges from, other languages within the language family it belongs to.<sup>112</sup> In the former case, the word is not so much considered as an object of use in communication, but rather as a phenomenon liable to etymological explanation. Analogously, in the latter case, the language is regarded as a target of explanation in historical or comparative linguistics, rather than as a system of symbols. In general, symbols or symbol systems are objects of factual understanding insofar as attention is being paid to the non-symbolic categories those symbols and symbol systems *also* belong to – next to, or by virtue of, them being symbols or symbol systems.

Two further remarks about my way of delineating factual understanding are in order. First, and perhaps surprisingly, the demarcation criterion implies that *meanings* are targeted by factual, rather than by symbolic understanding. To

---

<sup>112</sup> I grant that it is not very common to express such understandings in terms of the construction ‘to understand + a word/a language’. But the use of this construction for the purpose of expressing factual understanding does not strike me as necessarily infelicitous – it might just require cancelling some Gricean implicatures.



#### 4 Factual understanding

understand a linguistic expression is to have symbolic understanding, but to understand the meaning of that expression is to understand something which is not symbolic: meanings are facts or phenomena, not symbols.<sup>113</sup> By the same token, facts about the compositional nature of a symbol are also targeted by factual understanding. When one understands what a sentence's grammatical subject, verb and direct object are, one has factual understanding about (part of) the sentence's internal structure. Since appraisal of such 'meta-symbolic' understanding sometimes features in everyday assessments of linguistic understanding, it may intuitively be associated with (a pretheoretic notion of) symbolic understanding. Here, however, the cited demarcation criterion firmly relegates all forms of meta-symbolic understanding to the domain of factual understanding.

The second remark is meant to resolve an issue mentioned at the end of section 1.1. There, I pointed out that it may sometimes be difficult to label certain targets of objectual understanding as either factual or symbolic. For instance, is understanding a narrative a matter of factual, or of symbolic understanding? The demarcation criterion allows us to answer this. If the narrative is simply conceived of as a sequence of symbols (e.g. sentences), then the appropriate epistemic category is symbolic understanding. This is admittedly a 'lower' form of narrational understanding: it merely involves understanding the symbols of which the narrative is composed. By contrast, if the narrative is thought of as something which can be characterized in terms of claims about a fictional world (e.g. 'in the narrative,  $x$  is the case') or in terms of claims about the actual world ('the narrative provides insight into real-world phenomenon  $y$ '), then the relevant category is factual understanding. Something similar goes for understanding concepts. Such understanding falls with the purview of symbolic understanding in case concepts are taken to be linguistic entities. When, on the other hand, concepts are thought of as psychological entities or as abstract objects, the relevant category will be factual understanding.

Now that symbolic and factual understanding have been distinguished in more precise terms, it is apt to start my inquiry into the latter concept with a concise overview of the contemporary epistemological literature on the topic. Factual understanding has, often under different terminological guises, generally been conceived of in cognitive terms. On the most traditional version of this conception, the unit of factual understanding is taken to be *belief*. This view has it that a subject understands a phenomenon or a subject matter when she has beliefs about it that meet a *veridicality condition* such as truth or truth approximation (De Regt & Gijsbers, 2017, p. 50). That veridicality condition is meant to secure a link between the

---

<sup>113</sup> Then again, to understand a sentence which expresses the meaning of a sentence is, once more, to have symbolic understanding. Assuming that meanings can be understood in terms of understanding sentences which express those meanings, it follows that one can in principle simultaneously possess (1) symbolic understanding of a sentence, as well as (2) factual understanding of the meaning of that sentence through (3) having symbolic understanding of a sentence which expresses that meaning.

subject's beliefs and the external world, thus answering to the intuition that having factual understanding should imply having genuine access to our mind-independent environment.

It is not difficult to see how a traditional, belief-based conception of factual understanding, conjoined with a veridicality requirement, could prompt one to endorse reductivism about understanding: the view that understanding equates ordinary propositional knowledge. Sliwa (2015), for example, partly commits herself to such a view by defending a reductivist account of what may be called 'interrogative understanding' – the form of 'to understand ...' that takes an interrogative clause as its complement ('to understand + why/what/which/when').<sup>114</sup> Other authors have endorsed more qualified forms of reductivism about variously labelled forms of understanding that coincide or overlap with my concept of factual understanding. Kelp (2017), who focusses on understanding phenomena (a proper subset of factual understanding), argues that such understanding is ultimately a matter of knowing propositions, but adds that we need to couple this view with some metric that accounts for understanding's inherently scalar character.

Moving away from reductivist accounts of factual understanding, we encounter a sizeable contingent of authors who agree that factual understanding is cognitive, but who deny that the unit of factual understanding is belief. A notable proponent of such a view is Grimm (2014), whose interest lies with explanatory understanding – another category that overlaps with factual understanding. Grimm claims that while explanatory understanding equates knowledge of causes, such knowledge is different from ordinary propositional knowledge. As he writes, the kind of knowledge of causes relevant to explanatory understanding crucially involves "seeing or grasping the modal relatedness of the terms of the causal relata" (Grimm, 2014, p. 336). In other words, knowledge of causes has as its object not propositions, but rather modal relationships. In speaking of *grasping*, Grimm sides with other voices in the understanding debate who have promoted the idea that it is this concept, rather than belief, that should feature centrally in a definition of factual understanding. Thus far, the concept of grasping has remained somewhat undertheorized. In an early contribution to the understanding literature, Riggs (2003, p. 20) rests content with the elusive claim that understanding a subject matter "requires a deep appreciation, grasp or awareness of how its parts fit together, what role each one plays in the context of the whole, and of the role it plays in the larger scheme of things". Even more non-committally, Strevens (2017, p. 41) writes: "To grasp a fact is like knowing the fact, but it involves a more intimate epistemic acquaintance with the state of affairs in question."

---

<sup>114</sup> That she is so committed derives from the fact that on an epistemological level, factual understanding has considerable overlap with interrogative understanding: what is expressed by 'to understand + a noun (phrase)' can often be expressed without loss of meaning in terms of the grammatical construction 'to understand + wh...'. See also section 1.1 for some more examples of overlap among different types of understanding.

#### 4 Factual understanding

Sketchy as these claims are, the notion of grasping has nonetheless gained traction among epistemologists and philosophers of science. Since grasping, unlike belief, is intuitively associated with knowledge-how rather than with knowledge-that, the understanding debate has seen a gradual shift towards defenses of ability-oriented accounts of factual understanding. This, in turn, has propelled the philosophical study of the epistemic role of behavior, although it must be added that overtly behaviorist theories of factual understanding are still few and far between.<sup>115</sup> In addition to this ‘reluctant behavioral turn’, there is a growing sense that factual understanding cannot be reconciled with an unrelenting demand for veridicality. Citing many examples from the exact sciences, proponents of non-factive accounts of understanding such as Elgin (2017) and Potochnik (2017) have shown that in employing models, scientists inevitably rely on idealization, and that the veridicality-violations involved in idealization are often epistemically benign. In fact, both authors argue that idealizations often provide us with epistemic goods that more veridical representations may not be able to deliver. According to Elgin, idealizations succeed in rendering salient features of reality that would otherwise elude us. In a similar vein, Potochnik (2020, p. 940), writes:

[I]n many (...) instances, the path to scientific understanding is paved with falsehoods, that is, with idealizations. This is so whenever representing an illuminating causal pattern is benefited by setting aside complicating details, details that may be causally relevant in their own right but that are incidental to the pattern focal to immediate research. It is in this way that idealizations, falsehoods, can directly facilitate understanding.

In section 4.5, I return to the topic of idealization when I discuss the benefits of my own proposed account of factual understanding.

By way of summarizing the foregoing, we can discern two trends within the contemporary debate about factual understanding. First, although the large majority of epistemologists still subscribes to a cognitive conception of factual understanding, there has been a growing interest in the epistemic significance of behavior, signaled by the advent of the notion of grasping. And secondly, increased focus on the subject of idealization has occasioned the reconsideration of the importance of veridicality to factual understanding.

I shall now elaborate on my own preferred approach to theorizing about factual understanding, and thereafter explain how that approach connects to existing views.

---

<sup>115</sup> Perhaps more so than any other recent contribution to the understanding debate, De Regt’s influential book *Understanding Scientific Understanding* (2017) has been instrumental to this development. In that book, De Regt argues that in order to understand phenomena, scientists need to construct models of those phenomena on the basis of intelligible theories, hence implying that behavioral (and not just cognitive) abilities need to be included in a theory of understanding.

#### 4.1 Towards an explication of factual understanding

Recall that in section 1.5, I formulated the following explicandum of factual understanding:

*Explicandum of factual understanding*

A subject  $S_1$  has a better understanding of some factual target  $\Psi$  compared to subject  $S_2$  if and only if compared to  $S_2$ ,  $S_1$  (ii) has a disposition to form superior cognitive states concerning  $\Psi$ .

As said, most authors have stayed relatively close to (a non-comparative version of) this pretheoretic conception of factual understanding. Given the widespread adherence to the intuition-based method of conceptual analysis, this is unsurprising. My aim, however, is to transform the pretheoretic conception of factual understanding into one that is tailored to the function I take ascriptions of factual understanding to have.<sup>116</sup> I thus intend to replicate the approach previously taken to explicating symbolic understanding.

I argued in chapter 2 that the aim of ascribing symbolic understanding consists in preserving stability in the use of symbols and interpretative rules with an eye to maintaining communicative efficiency. My argument for this claim was based on the observation that a community simply *needs* such stability, given that its proper functioning as a community would otherwise be seriously jeopardized. I also showed that from this vantage point, ascriptions of symbolic understanding should be regarded as pertaining to behavioral dispositions. The question now is: what would this kind of approach entail for how we explicate factual understanding? Citing again from section 1.5, ascriptions of factual understanding serve the following purpose:

*The function of ascriptions of factual understanding*

An ascription of factual understanding to an individual serves to encourage that individual to sustain the possession of a disposition that contributes to optimizing a community's attunement to its environment.

By attunement I mean the following: a community is more attuned to its environment the more the community at large, and individuals and groups within the community, are able to achieve their goals through engaging in efforts to achieve those goals.<sup>117</sup>

---

<sup>116</sup> One could object that there is in fact no practice of ascribing factual understanding. After all, we do not usually say things like 'A factually understands  $x$ ', or 'B has factual understanding of  $y$ '. This objection misses the mark, however. I do not claim that there exists a practice in which the term 'factual understanding' (or any of its cognates) is literally applied. All I am saying is that it is an empirical fact that people use expressions of the type 'to understand + noun (phrase)', and that we can discern within the general category of such ascriptions two practices which have distinct purposes. I have chosen to refer to these practices by using the terms 'symbolic' and 'factual understanding'.

<sup>117</sup> Note that this means that achieving one's purpose through sheer fortuity does not positively impact on one's attunement to the environment.

#### 4 Factual understanding

In order for such efforts to stand a chance of success, agents must be able to anticipate the effects of those efforts. It is this ability to anticipate effects which the practice of ascribing factual understanding aims to foster. Thus, where ascriptions of symbolic understanding serve to maintain communicative efficiency, ascriptions of factual understanding serve to ensure that agents can reliably anticipate the effects of encounters with their environment (which may include the behavior of other agents).<sup>118</sup>

Just as the practice of ascribing symbolic understanding involves a means through which its main purpose – maintaining communicative efficiency – is achieved, so does the practice of ascribing factual understanding. In the former case, the means is the furtherance of symbol-usage that accords with established rules of interpretation. In the latter case, the means is the furtherance of informational transmission. More specifically, the purpose of ascriptions of factual understanding is to promote the possession of dispositions for the dissemination of certain information; namely information by means of which members of the community can foresee the effects of worldly events and of their own actions, and thereby achieve their goals more effectively. The factual understander is an agent who can be relied upon to provide others with epistemic goods that enable those others to anticipate whatever may happen in the world around them. Seen in this way, the practice of ascribing factual understanding caters to a fundamental human need. For given that agents within a group are bound to be unequally privileged as to their access to relevant environmental information, there will need to be a continuous exchange of such information in order to make sure that each individual in the group can adequately perform the role upon the execution of which his fellow group-members rely.

Admittedly, this is not the only way in which one could think of the purpose of ascribing factual understanding. A perhaps more obvious alternative would be to say that factual understanding is ascribed in order to promote the acquisition of true beliefs and the elimination of false ones, where truth is conceived of in a standard, correspondence-to-reality kind of way. The question is, however, whether communities have a genuine *need* for such truth-based attunement. In order for a community to function as a community, its members must succeed in the pursuit of their goals sufficiently frequently. After all, it would be difficult to imagine a scenario in which the day-to-day activities of a group of agents (including the most trivial) were consistently fruitless and misdirected, but in which the group would still recognizably operate as a community. Simply put, then, things generally ought to *work*: agricultural activities should typically procure food, medical interventions should reliably restore and enhance health, weather forecasts should frequently yield predictions that match actual events, etc. Clearly, therefore, it is in the interest of a

---

<sup>118</sup> Encounters with the environment include confrontations with behavior of environmental targets over which one has no or little control (e.g. weather events, asteroids), as well as with the results of manipulations and interventions (kicking a ball, conducting a scientific experiment).

#### 4.1 Towards an explication of factual understanding

community to promote behavior that increases the likelihood that efforts at attaining goals are successful.

But should, in addition to this, communities want to encourage biologists, medical scientists and meteorologists to aspire to truth-as-correspondence? This is far from evident. Perhaps a community needs truth in a strategic sense. That is to say, the desire for truth, and the implicit endorsement of a particular conception of truth, may turn out to be necessary for motivating agents to put in the effort that is needed for things to work. But this does not imply that communities need to promote truth as such. A community needs to be successful in certain regards in order to function properly as a community, and this just means that the community, and thereby its members, must have a strong enough tendency to engage in fruitful activities. From this perspective, truth is not a genuine need of a community. By implication, there is no need within a community to foster attunement in an *alethic* (=truth-based) sense. It only needs to foster attunement in a *practical* sense. My conception of the function of ascribing factual understanding is predicated on this latter sense of attunement, based on the idea that we should not multiply needs beyond necessity.

It is worth highlighting two possible objections to this line of thought. First, one may point out that the acquisition of true beliefs and the uttering of true assertions is itself among the things that may or may not work. On this view, acquiring a true belief, or uttering a true assertion, is a purpose, the pursuit of which may fail or succeed. If it fails, this impacts negatively on the community's proper functioning. In section 4.4 (#4), however, I will argue – or rather, repeat an argument made by Peirce – that aspiring to truth-as-correspondence is not a genuine purpose, since we have no means for discriminating truth from non-truth. For we ultimately do not know what a belief or assertion has to be like for it to be true: arriving at the truth is not a type of outcome which we can readily distinguish from arriving at falsehood.

A second objection to my claim that truth-as-correspondence is not a genuine need of a community, is that a pragmatist conception of attunement may presuppose truth indirectly, namely in the sense that things (e.g. agricultural activities, medical interventions, weather forecasts) are only likely to work if they are based on true beliefs and assertions. After all, what would explain a community's various successes if not their being grounded in truth? This reminds us, of course, of the well-known 'no-miracles argument' for scientific realism.<sup>119</sup> I will not take sides within the long-standing debate over this argument, but only point out that even if this kind of argument succeeds, a pragmatist conception of attunement remains the preferred option, for two reasons. First, even if truth were required for things to work, the acquisition of true beliefs and the uttering of true statements could only ever be promoted *via* the promotion of the dissemination of useful information. For given that truth is not a genuine purpose (see, again, section 4.4), truth can only be strived

---

<sup>119</sup> Putnam (1975, p. 73) is usually cited as the original source of this argument.

#### 4 Factual understanding

for indirectly – namely via the pursuit of purposes that *are* genuine. Secondly, even if truth were a genuine purpose, a community should not want to prioritize the acquisition of true beliefs, or the uttering of true statements, over the dissemination of useful information. After all, not all truths are equally relevant to a community's proper functioning. If a community were to indiscriminately encourage the pursuit of truth (granting this is possible), its proper functioning would be safeguarded relatively inefficiently compared to a community that encouraged the dissemination of useful information.<sup>120</sup>

My contention, then, remains that communities should be concerned with practical attunement, and hence that they should want to encourage the possession of dispositions that contribute to such attunement. By assumption, ascriptions of factual understanding cater to this demand. An important implication of this is that factual understanding consists in *behavioral* dispositions. For I have argued that to contribute to a community's environmental attunement is to disseminate useful information, and to disseminate useful information is to display behavior. My arguments against the relevance of cognitive dispositions to factual understanding mirror those I put forward in defense of my behaviorist conception of symbolic understanding. To cut a long story short: since cognitive dispositions are only of indirect importance to fostering environmental attunement, they do not constitute a unit of epistemic assessment. A community needs to foster specific behavioral dispositions, but it does not need to independently foster specific cognitive dispositions. In fact, in keeping with the arguments from section 3.1, a community should want to allow for a wide spectrum of cognitive dispositions in optimizing its members' dispositions for informational transmission. Hence, there is ample reason for a community to *not* have its concept of factual understanding apply directly to cognitive dispositions.

In closing, let me comment on how my preferred account of factual understanding connects to the two contemporary trends identified previously: the reluctant behavioral turn and the reconsideration of the importance of veridicality. As to the first trend, I am allying myself to what is (still) a minority position within the current debate. To be sure, my aim in doing so is not to offer a novel account of grasping. Rather, my account of factual understanding ties in with the reluctant behavioral turn more or less incidentally, by virtue of its aim to capture not our intuitions, but instead the *raison-d'être* of a certain practice. My proposal aligns with the second trend too, in the sense that it does away with the all too rigid truth condition on factual

---

<sup>120</sup> It is worth noting that a similar argument can be raised against the idea that communities should promote the acquisition of truth-as-coherence. For although aspiring to coherence with existing beliefs could be regarded as a genuine purpose (given that we can in principle verify whether something does or does not cohere with our existing beliefs), coherence in itself does not necessarily safeguard a community's proper functioning. For whereas it may be plausible to say that properly functioning communities will tend to have coherent belief systems, coherent belief systems do not entail proper functioning.

#### 4.1 Towards an explication of factual understanding

understanding. Instead of an alethic requirement, it imposes a pragmatic one: factual understanding demands not that one possesses true beliefs or utters true statements, but rather that one contributes, via the dissemination of information, to the achievement of purposes.

Although there is thus some continuity between my approach and existent theories of factual understanding, the above picture nonetheless implies a rather radical departure from the epistemological tradition. For I take it that few would be willing to stray so far from our intuitions as to say that factual understanding is exclusively a matter of behavior, let alone a matter of information-transmitting behavior. To legitimize my drastic Carnapian move, I need to develop the account of factual understanding suggested above into a more complete theory. Once that is done, I must show that this theory has various benefits over and above its ability to explain why we have a practice of ascribing factual understanding. Furthermore, in order to let those benefits accrue to the explicatum of symbolic understanding, I have to make clear that symbolic understanding, as explicated by me, figures ineliminably in the explicatum of factual understanding. In the next section, I take on the first and third of these tasks, leaving the second to be dealt with in section 4.5.

#### 4.2 Explicating factual understanding

In the previous section, I outlined my preferred account of factual understanding. Few details have been provided yet, however. Furthermore, I have not yet shown how my account of factual understanding hinges on the explicatum of symbolic understanding I have formulated in the previous chapter. Doing so would confer the much-needed, additional justificatory support upon that explicatum.<sup>121</sup> My task in this section is thus to explain how factual understanding can be explicated using the proposed explicatum of symbolic understanding. I will argue that since factual understanding is ascribed on the basis of an agent's disposition for transmitting useful environmental information, symbolic understanding is indispensable – indeed part and parcel of – factual understanding. The end result of this section will be a definition of factual understanding. Although I aim to be as comprehensive as possible, some details will be left unspecified. Where details are omitted, I shall hint at directions for further specification. As a further disclaimer, I should add that in light of the ground that is covered in this section, I am compelled to postpone the provision of clarifications, the discussion of possible objections, and the formulation of positive arguments to sections 4.3, 4.4 and 4.5, respectively.

What exactly does it take to understand phenomena, historical events, persons, or, indeed, meanings? I explained earlier that in order to answer this question, we should look at how the practice of ascribing factual understanding is meant to benefit

---

<sup>121</sup> I say 'further' justificatory support, because some of that support was already meant to be procured by the fact that my account of symbolic understanding (partially) satisfies Carnap's similarity, simplicity and exactness conditions.



#### 4 Factual understanding

the community of agents within which that practice exists. My suggestion was that factual understanding is best seen as a tool used for optimizing the environmental attunement of a community of mutually dependent agents. The factual understander is someone who is disposed to enable others to successfully anticipate (the effects of) their encounters with the world, so that those others can better fulfil the various roles upon the execution of which fellow community-members rely. Factual understanding is thus an intrinsically social affair, in the sense that it consists in dispositions for other-directed behavior.<sup>122</sup>

I will now argue that if we start from this general picture, factual understanding *always* requires symbolic understanding. To see this, the first thing to note is that disseminating information is necessarily purposeful behavior. For instance, if you are approached by my dog upon entering my property, and I tell you that my dog barks out of enthusiasm rather than aggression, I do so to achieve some purpose. Most probably, that purpose is to let the information expressed by my utterance aid you in deciding how to approach the dog. In this case, the means of transmission is recognizably symbolic: the telling, we may assume, happens through the use of sentences of a natural language. But we can also think of other ways of transmitting environmental information which do not, at first sight, involve symbols. For example, a driving instructor may demonstrate how and when to change a car's gear without any verbal elucidation, through letting the car's responses to her interventions (e.g. its sounds and speed changes) do the educational work. Such a demonstration may assist the aspiring driver in his own driving attempts, thus facilitating the anticipation of future encounters with his environment.

On further thought, however, the second case of informational transmission is as much symbolic as the former. The communicative purpose of the driving instructor makes it so. For in manipulating the clutch and the gear lever, and in letting the car produce the varyingly disconcerting effects associated with mistimed gear changes, the driving instructor aims to let her actions and the effects of those actions be interpreted in a certain way. That is, the driving instructor purports to have her apprentice regard (certain aspects of) the instructor's actions, together with their antecedent conditions and consequences, as signs of ideal and non-ideal procedures. Per what was said in section 2.5, such signs are symbols. To be sure, the driving instructor's non-verbal instructions function in a way that is somewhat different from how conventional symbols function. The rules of interpretation that govern linguistic expressions, for instance, ordinarily obtain prior to any language-user's use of those expressions. This means that the proper use of a natural language expression is

---

<sup>122</sup> This is not to say, however, that such behavior need be altruistic. For purposes to provide others with epistemic goods can be nested within further, purely self-centered purposes. An agent's endeavor to transmit environmental information to another agent can have as its ulterior purpose to buttress the social position of the former agent, or to exercise control over the latter agent. The ascription of factual understanding does not see to the sanctioning of such ulterior purposes; for that aim we have ethical concepts.

rightly seen as a form of compliance: namely compliance with a purpose with which that expression was invested prior to, and independently from, any contemporary individual language-user's use of it. The driving instructor, on the other hand, is not so much complying with a rule of interpretation, as she is creating one. In acting out her non-verbal instructions, she uses a newly created symbol in accordance with the significatory purpose it has in that same act been invested with. There is thus propriety, but no propriety-as-compliance.<sup>123</sup>

The main takeaway from the above is that the transmission of environmental information, given that it necessarily involves communicative purposes, invariably requires the use of symbols. Indeed, it requires the *proper* use of symbols: the use of symbols for their significatory purpose. But does it also require *competent* symbol-usage, and therefore, as I set out to argue, symbolic understanding? My answer is yes, an answer that I shall now try to substantiate. As a prefatory step, consider first of all the fact that factual understanding, like symbolic understanding, must come in the form of dispositions. In order for a community to function stably over time, candidate-understanders are to be judged based not on their actual, but on their possible performance. We want our epistemic informants not just to provide us with useful information here and now, but to be disposed to do so across a range of possible circumstances. So factual understanding is dispositional. This, in turn, entails that our epistemic informants must be disposed to properly use the symbols needed for conveying the relevant environmental information. So factual understanding, given that it pertains to dispositions, requires *dispositional proper symbol-usage*.

This, however, does not yet establish that factual understanding requires *competent symbol-usage*. After all, dispositional proper symbol-usage can in principle stem from sources such as luck, mere memorization and mere imitation (see section 3.3). Agents who are disposed to reliably provide us with environmental information through the proper use of symbols, need not have acquired that disposition through learning (in the sense of section 3.4). Imagine a parrot successfully directing guests in a restaurant to the rest room by shrieking "toilets to your left!" every time a human being walks in its direction. As the parrot thus has a disposition to enable others to navigate through their environment, it may be deemed to possess factual understanding. But the parrot does not possess symbolic understanding of the expression 'toilets to your left', if we assume (as is plausible) that its disposition for using that expression properly was acquired on the basis of mere imitation and memorization. So it seems factual understanding does not

---

<sup>123</sup> There is arguably still some sense in which the driving instructor must also heed certain communicative conventions that were already in effect. For her symbols can only operate under a system of tacit assumptions and mutual expectations associated with driving lessons and didactic contexts more generally. If the non-verbal instructions of the kind the driving instructor is attempting to convey were an altogether alien category within such contexts, her efforts would likely be in vain.

#### 4 Factual understanding

necessitate symbolic understanding. Here, however, I want to rekindle an argument used earlier in the context of rejecting luck, imitation and memorization as possible sources of competence. As said in section 3.3, those sources ultimately cannot sustain a system of efficient communication. Yet such a system is needed if a community is to have a durable practice of exchanging environmental information; it cannot ground such a practice on an unsustainable communicative system. So although we can in principle conceive of *instances* of factual understanding based upon non-competent proper symbol-usage, we cannot imagine a temporally extended *practice* of ascribing factual understanding based upon such symbol-usage. Hence factual understanding requires symbolic understanding after all.

Based on the above, we can formulate the following necessary condition on factual understanding:

*Necessary condition #1 on factual understanding*

An agent must be disposed to use some symbol  $\Phi$  competently.

Clearly, symbolic understanding is not by itself sufficient for factual understanding: one can use symbols competently without aiming to inform others about their environment. Some types of symbols do not even lend themselves to being put to such a purpose: imperatives and interjections are a case in point. So, in order to arrive at a satisfactory account of factual understanding, we need to add some further criteria. In what follows, I will formulate two such criteria which, together with the requirement of symbolic understanding, I take to be jointly sufficient for having *some (comparative) degree* of factual understanding. This ‘blueprint’-definition of factual understanding will ultimately have to be complemented with an account of how factual understanding can vary in (comparative) degree. However, since my aim here is to show how the explicatum of symbolic understanding can be fruitful for developing an account of factual understanding, I take a blueprint-definition to constitute a satisfactory result within the context of this investigation. Nonetheless, at the end of this section, I will offer some hints as to how degrees of factual understanding might be accounted for.

I said earlier that factual understanding is ascribed to those agents who are disposed to enable other agents to anticipate (the effects of) encounters with their environment. For an agent to be a reliable informant of this kind, he must be disposed to manifest his disposition for competent symbol-usage *in order to* provide others with information useful for the achievement of some purpose. Hence the following:

*Necessary condition #2 on factual understanding*

An agent (the sender) must be disposed to manifest his disposition for using symbol  $\Phi$  competently for the purpose of (i) transmitting information about some factual target  $\Psi$  to another agent (the recipient), and (ii) thereby enabling

the recipient to achieve a certain purpose pertaining to  $\Psi$ , which the sender believes, conjectures, or presupposes the recipient has or could have.<sup>124</sup>

(i) and (ii), although separately indicated, are in fact inseparable: it is by virtue of (ii) that the act of using a symbol qualifies as serving purpose (i). That is to say; a symbol becomes a carrier of information through its being used in order to enable the achievement of certain goals. Indicating (i) and (ii) separately, however, allows me to state condition #2 more clearly: it should be the sender's purpose to enable the recipient to achieve certain purposes *by means of* the informational transmission manifested through the sender's symbol-usage.

As to the theoretical implications of condition #2, the final part of clause (ii) is especially important. For it entails that the attribution of factual understanding is sender-centered.<sup>125</sup> It is the purpose the sender thinks the recipient has or could have, not the purpose the recipient actually has, which matters to factual understanding. The rationale behind this is that a recipient-centered view would be too demanding in light of what the concept of factual understanding exists for.<sup>126</sup> It would, for instance, deny factual understanding to the agent who is disposed to provide others with potentially highly useful information about, let's say, the fundamental nature of the universe, if it were the case that those others had no interest in that information. When it comes to optimizing a community's attunement to its environment, it is important that agents are maximally disposed to share information pertaining to that environment. If the sender can think of some possible  $\Psi$ -related purpose, however trivial or far-fetched, to which his information is relevant, then he should be rewarded for contributing to the achievement of that purpose via the dissemination of the information.

The criterion just formulated does not yet complete the picture. After all, it is not enough for a sender to aim to benefit the recipient through transmitting information;

---

<sup>124</sup> Although I will not press the point further here, I think we should, at least in the present context, think of beliefs, conjectures and suppositions as dispositions rather than as occurrences. On this conception, the sender need not actively entertain the belief, conjecture or supposition that the recipient has, or could have, a certain purpose. Provided that we construe the stimulus conditions of the sender's dispositional mental state in an appropriate way, condition #2 can thus accommodate the not improbable scenario in which the sender's mental state is not fully transparent to him even whilst he engages in the transmission of information.

<sup>125</sup> In order to defuse a potential worry: this does not contradict my earlier claim that factual understanding is other-directed. Factual understanding is other-directed in the sense that it consists in dispositional behavior which is intrinsically geared towards other agents. This is reconcilable with the idea that normative assessments of such dispositional behavior are attuned to the perspective of the sender.

<sup>126</sup> So, to be clear, I am not claiming that a recipient-centered view would be *intuitively* too demanding. In fact, a recipient-centered view may well be more intuitively plausible than a sender-centered view. But, as repeatedly stated, maximizing conformity to our intuitive judgments is not what my approach is aimed at.

#### 4 Factual understanding

it must also be the case that this aim is attained. A horoscopist may aim to enable others to predict the influences of planetary alignments on earthly events by transmitting information about those purported influences, but most would agree that he cannot really achieve this aim.<sup>127</sup> It is only when a sender's transmittory purpose is achieved, that his transmission can be said to contribute to the goal of optimizing a community's attunement to its environment. For this reason, a community should want to deny factual understanding to astrologers and other sources of less than useful information. To take this into account, I propose the following, third criterion on factual understanding:

*Necessary condition #3 on factual understanding*

When a sender transmits information in accordance with condition #2, the sender's purpose to benefit the recipient must be achieved, and the achievement of that purpose must be due to the transmission.

This condition can easily be misunderstood, so I must clarify it briefly. By the requirement that the purpose of the sender be achieved, I do not mean that the recipient must benefit from the transmission in the sense of actually achieving the purpose the sender attributes to her. This, again, would be too strict a demand. For first of all, as said before, the recipient may not (currently) have the purpose attributed to her. And secondly, even if the recipient does have that purpose, she may fail to monetize on the transmission for reasons which have nothing to do with the usefulness of the transmitted information, or with the manner of transmission. After all, she may fail to realize the sender's intended purpose solely because she herself errs avoidably in her attempt to achieve it. Condition #3 imposes a 'softer' demand instead. It says, per its reference to condition #2, that the transmission must *enable* the recipient to achieve the purpose attributed to her. This means that it is, strictly speaking, not the purpose of the sender to have the recipient actually achieve the attributed purpose (although this would obviously be a desirable effect of the sender's transmission), but rather to put her in a position to achieve that purpose. In other words, for condition #3 to be met it must be the case that the transmission puts the recipient in a position to achieve a purpose which the sender attributes to her.

The three criteria just formulated jointly make for a blueprint-definition of factual understanding. That is, in case an agent satisfies all three conditions, she possesses

---

<sup>127</sup> Of course, a horoscopist can succeed in other respects. He may enable the recipient to form a world picture, or personal philosophy, which can improve the recipient's psychological well-being. But insofar as this is a purpose of the horoscopist, it is not a purpose that pertains to the influence of planetary alignments on earthly events. If it were the horoscopist's explicit purpose to transmit information about the recipient's psychological susceptibility to astrological musings through engaging in those musings, however, his success in improving the recipient's psychological well-being might indeed warrant an ascription of factual understanding. But then it would be factual understanding of a different  $\Psi$ : namely the recipient's psychological nature.

some (comparative) degree – this can be any degree – of factual understanding. As said, further work is needed to spell out how those degrees are determined. Below, I will give some hints in this direction. But let me first amalgamate the three criteria just formulated, with some notational adjustments, as follows:

*Factual understanding*

An agent  $A$  possesses some degree of factual understanding of some target of factual understanding  $\Psi$  if and only if

- (i)  $A$  has a disposition  $D_\phi$  for using some symbol  $\Phi$  competently, and
- (ii)  $A$  has a disposition  $D_\psi$  to manifest  $D_\phi$  for the purpose of (a) transmitting information concerning  $\Psi$  to some other agent  $B$  and (b) thereby putting  $B$  in a position to achieve a certain purpose pertaining to  $\Psi$  which  $A$  believes, conjectures or presupposes  $B$  has or could have, and
- (iii) If  $D_\psi$  is manifested, then  $A$ 's purpose to benefit  $B$  in the way cited under (ii) must be achieved, and that achievement must be due to the manifestation of  $D_\psi$ .

An example may help to illustrate this definition. What does it take, for instance, to have some degree of understanding of the phenomenon of solar eclipses? The definition in principle allows for innumerable scenarios. I will highlight just one possibility.

First, according to (i), there must be some symbol which an agent (the sender) is disposed to use competently. Let this symbol be the sentence 'Solar eclipses are caused by the Moon's obscuring the Sun's light from the perspective of an observer on Earth'. Condition (ii) says that the sender must be disposed to use this sentence for the purpose of transmitting information regarding the phenomenon of solar eclipses to a recipient, such that the recipient is thereby put in a position to achieve some purpose (attributed to the recipient by the sender) having to do with that phenomenon. Assume that the purpose the sender attributes to the recipient is to (roughly) determine under which sorts of conditions solar eclipses are liable to occur. Criterion (iii), finally, tells us that the sender's purpose to enable the recipient to achieve the purpose attributed to her, must be achieved if the sender actually engages in said transmission. So her uttering the above sentence must put the recipient in a position to roughly determine under which sorts of conditions solar eclipses are liable to occur. Let us assume this to be the case: the information the sender provides indeed puts the recipient in a position to infer that solar eclipses only occur when the Moon is positioned in between the Sun and the Earth – an inference which is confirmed by ample historical evidence and which we may presume to be further

#### 4 Factual understanding

corroborated by future evidence. By virtue of this, the sender has some degree of understanding of solar eclipses.<sup>128</sup>

Factual understanding, like symbolic understanding, is intrinsically scalar. So the above definition leaves a lot to be desired: we would ideally like to know what makes for comparatively better or worse factual understanding. Without going into much detail, here are some parameters that I think should figure in an account of (comparative) degrees of factual understanding:

- *Breadth*. For how many, and what sorts of, symbols does a sender satisfy conditions (i-iii)? For instance, is the sender able to report only on one aspect of  $\Psi$ , or can she also report on many others?
- *Depth*. For the achievement of how many, and what sorts of, different purposes relating to  $\Psi$  can a sender provide useful information? For instance, is the sender capable of enabling only the inferring of general claims about  $\Psi$ , or can she also put her recipients in a position to make highly accurate predictions regarding  $\Psi$ ?

---

<sup>128</sup> What about non-human animals? Can they possess factual understanding as well? This question can be approached in two ways, one uninteresting, one considerably more interesting. The uninteresting approach says that factual understanding, being a human language term, only has application within the human sphere. As also explained in footnote 41, however, this is not really what we want to say. What we would like to know is whether there are non-human animals which, were they to have an evaluative instrument which functioned in all the same ways that our concept of factual understanding functions (i.e. as a tool for promoting dispositions for informational transmission), could qualify for ascriptions of their form of factual understanding. I do not possess sufficient ethological expertise to answer this, but it is worth pointing to one issue that would have to be addressed. An important fact of my notion of (human) factual understanding is its fundamental other-directedness: ascriptions of factual understanding pertain to dispositions for enabling others to achieve purposes. As Tomasello (2008, p. 52-53) observes, the human capacity for other-directed action is remarkable in that it is not encountered to nearly the same extent in other hominids:

When one human points for another, the recipient implicitly asks herself why— why does he think that looking in that direction will be useful or interesting for me? This is based on the assumption that he is indeed pointing for her benefit (at least immediately). Thus, young children know that an adult's pointing to a bucket in the context of a searching activity is probably relevant in some way to their joint goal of finding the toy. In contrast, great apes cannot and do not assume that the other is pointing for their benefit, and so they do not ask themselves "why does he think this is relevant for me?" They want to know what he wants for himself (since when they point it is always for themselves), not how he thinks their looking in this direction will be relevant for them— and so they simply do not see another's pointing gesture as relevant to their own goal.

I am reluctant to generalize from this example, but I think it provides some reason to believe that our capacity for being credited with factual understanding may be a property that is relatively hard to come by elsewhere in the biosphere.

- *Robustness*. How comprehensive is the set of stimulus conditions associated with a particular disposition  $D_\psi$ ? For instance, does the disposition only pertain to possible circumstances where the sender is able to rely on external sources of information (books, experts), or does it generalize to possible circumstances in which she has only her own mind to assist her?
- *Specificity*. How well is the sender's transmission tailored to the purpose attributed to recipient? For instance, can the recipient very easily attain the  $\Psi$ -related purpose based on the information the sender provides her with, or does she herself need to put in a significant cognitive effort too?

This list, which is by no means meant to be exhaustive, provides some indication of just how multi-faceted the gradual character of factual understanding is. I should add, however, that the above set of parameters needs to be complemented with a protocol for determining weights of purposes and of stimulus conditions (see also section 3.6). In the case of depth of factual understanding, for instance, it seems reasonable to say that not all transmittory purposes are equally significant. For one could argue that enabling others to make very accurate predictions is, generally speaking, a more weighty epistemic achievement than enabling others to only make very crude predictions. I would suggest that the epistemic weight of a purpose or a stimulus condition at least partly depends on how strongly the achievement of a purpose, or the manifestation of a disposition, contributes to the overall aim to which the concept of factual understanding is geared (i.e. optimizing a community's attunement to its environment).<sup>129</sup> Here, I will not attempt to construct a theory of epistemic weights around this idea, however.

---

<sup>129</sup> In section 4.5, #4, I point out that how a community conceives of the optimization of its environmental attunement may change over time. This entails that the way in which purposes are weighted may change as well. Comparative assessments of factual understanding are therefore not necessarily diachronically fixed. As I hinted at in section 3.6, something similar goes for comparative assessments of symbolic understanding: since rules of interpretation may (and sometimes need to) change, what counts as proper usage of a symbol at  $t_1$  need not amount to proper usage at  $t_2$ . As I also explained in footnote 108, however, this does not mean that ascriptions of symbolic understanding should be relativized to rules of interpretation, as that would too drastically undercut the effectiveness of such ascriptions for maintaining communicative efficiency. Neither, I think, should ascriptions of factual understanding be relativized to conceptions of optimal environmental attunement. For a community to function properly, there should be a considerable degree of internal agreement within the community as to which purposes matter most to its environmental attunement, and which less so, or not at all. This is because the vast majority of efforts at optimizing environmental attunement will be co-operative, in the sense that individual agents rely on one another in their attempts to attain their purposes. If ascriptions of factual understanding were relativized to conceptions of environmental attunement, those ascriptions would insufficiently promote the formation of the needed internal consensus, and thereby ultimately threaten the common ground that is needed for successful co-operative action. Over time, the consensus as to the nature of environmental attunement may shift. There is no need for



#### 4 Factual understanding

I also refrain from adding to the definition of factual understanding the kind of exemption clause for agents with (severe) behavioral limitations that was included in the explicatum of symbolic understanding. This is because I do not feel there is an equally strong rationale for doing so. In case of symbolic understanding, the admission of (a certain class of) counterfactual behavioral dispositions served to prevent the practice of ascribing symbolic understanding from becoming unduly exclusionary. That is, it was meant to rectify some of the inherent injustice in the fact that those who enjoy the privilege of being able to use symbols *ipso facto* determine the communicative norms that operate in a community. With the exemption clause, the underprivileged could continue to be recognized as participants in a communicative practice, and hence as members of a community. When it comes to factual understanding, however, I do not see an immediate need for a similar correction mechanism. It is true that those who are able to use symbols enjoy the privilege of being able to inform others about their environment. But they do not determine the usefulness of the information they distribute. So although there inevitably exist structural inequalities in terms of which agents are liable to ascriptions of factual understanding, those structural inequalities are not due to a ‘tyranny of the majority’ to the same extent that they are in case of symbolic understanding.

In this section I have explicated the concept of factual understanding, and in so doing, shown why and how symbolic understanding is relevant to factual understanding. The upshot is that factual understanding, when analyzed not on the basis of our intuitions but on the basis of the way the concept operates in a community, is a fundamentally intersubjective phenomenon. One has understanding of a factual target by virtue of having a disposition to successfully assist others in achieving their target-related goals through informing them about the target. Symbolic understanding is central to this conception of factual understanding for the following reason: since factual understanding implies informational transmission, informational transmission implies purposeful signification, and purposeful signification implies symbolization, factual understanding implies symbolic understanding. All this does not complete my treatment of factual understanding, however, as there is more to be said in defense of my proposal. For one thing, I have not yet clarified all of the potentially ambiguous components of the blueprint-definition. Secondly, I have not discussed any objections that may be raised against it. And last but not least, there are some further theoretical benefits of my account that are worth highlighting. These matters will be dealt with in the next sections.

---

communities to prevent such gradual, macro-level change by making sure that comparative assessments of factual understanding are diachronically fixed. For such gradual evolution does not hamper the proper functioning of communities at any given time. But communities *do* have to safeguard the internal consensus that is needed for joint action, by making sure that comparative assessments of factual understanding are synchronically fixed.

### 4.3 Clarifications

The proposal formulated in the previous section is undeniably a controversial one. First of all, it disavows the idea that factual understanding is cognitive. Secondly, it does away with truth as a necessary condition; replacing it with an ostensibly more permissive pragmatic requirement. And third, it says that symbolic understanding is a *conditio sine qua non* for factual understanding. Arguably, all these claims are at odds with our intuitions to at least some extent. While this need not in itself be problematic as long as the proposal is evaluated under the terms of its broader revisionistic commitments, there is still a strong onus on me to argue why my account is to be preferred over its more intuitively plausible contenders. For why be revisionistic if, in the end, the benefits do not outweigh the costs? To finalize my defense of the explicatum of factual understanding, I need to do three things. First, I must eliminate any potential sources of confusion lingering within the blueprint-definition. Second, I need to answer to the most pressing objections that may be levelled at that (clarified) definition. And third, I have to offer positive arguments for the claim that my proposal has considerable advantages over less controversial alternatives. The first of these tasks will have to precede the other two, given that my clarifications may bear on, and even give rise to, certain objections, and might in addition also impinge upon the benefits of my account. Three issues will be dealt with in this section: the interpretation of the phrases ‘concerning  $\Psi$ ’ and ‘pertaining to  $\Psi$ ’; the nature of the relation between sender and recipient; and the question what it means to put someone in a position to achieve a purpose.

#### 1. The meaning of ‘concerning $\Psi$ ’ and ‘pertaining to $\Psi$ ’.

The blueprint-definition does not impose any direct constraints on what symbols can be used for which targets  $\Psi$ . In principle, an agent can competently use any symbol to inform another agent about  $\Psi$ , and possess factual understanding of  $\Psi$  by virtue of this, as long as he satisfies criteria (ii) and (iii) as well. To be sure, for the use of some types of symbols, such as the imperatives and interjections mentioned earlier, it will be very difficult to satisfy those criteria. But such symbols are not excluded *a priori* – they are just highly unlikely to prove suitable *a posteriori*. In criterion (ii), two provisos are included to ensure that there is a link between an agent’s competent use of some  $\Phi$  and a target  $\Psi$ . The first proviso, cited under (a), is that agent  $A$  should have a purpose to transmit information *concerning*  $\Psi$ . And the second proviso, cited under (b), is that this transmission should enable an agent  $B$  to achieve a purpose *pertaining to*  $\Psi$ . The italicized phrases admit of what I will call internalist and externalist readings. In both cases, I opt for the externalist reading. Below, I substantiate this choice and subsequently comment on what I think determines whether something concerns, or pertains to,  $\Psi$ .

With respect to first proviso,  $A$  is free to use any symbol he prefers, but the information he aims to transmit in using the symbol (i.e. the information which is meant to enable  $B$  to do certain things), must concern, or be about,  $\Psi$ . This

#### 4 Factual understanding

requirement must be understood as pertaining to what is the case, and not to what *A* thinks is the case. That is, the information must *actually* concern  $\Psi$ . This means that when *A* has a purpose to transmit information, it is not required of *A* that he believes or assumes that this information concerns  $\Psi$ . Such an internalist reading would be too demanding, and potentially also too permissive at the same time. Too demanding, because if we take the practice of ascribing factual understanding to see to the effective anticipation of agents' encounters with their environment, there is no need for senders to be aware of some piece of information's relation to  $\Psi$ . For if we look solely at what is needed to enable recipients to achieve their (attributed)  $\Psi$ -related purposes, it is enough that senders transmit information which is sufficient for achieving this goal.<sup>130</sup> Still, an internalist reading of the first proviso might at the same time be too permissive. For it would mean, were we also to subscribe to an internalist reading of the second proviso, that *A* could have understanding of  $\Psi$  without there being an actual link between his information and  $\Psi$ . This would thwart the achievement of the general purpose of factual understanding: a community needs agents who *are* reliable informants, not just agents who think they are.

The second proviso, according to which the purpose attributed to *B* must pertain to  $\Psi$ , should also be interpreted externalistically. So while the attribution of a purpose to *B* is an affair internal to *A*, the (ir)relevance of that purpose to  $\Psi$  is something which is determined by facts external to *A*. The rationale behind this has already been hinted at above: the (potential) usefulness of *A*'s transmission with regard to  $\Psi$  must be grounded in reality, lest the practice of ascribing factual understanding become disconnected from that reality, and as a result, unsuitable for realizing the objectives of the practice itself.

The externalist readings of the two provisos oblige me to address a difficult question: what does it mean to say that something *actually* concerns, or pertains to,  $\Psi$ ? My answer to this would be that that depends on how a community chooses to conceive of  $\Psi$ . In fact, I would be inclined to say that the question which purposes a community deems relevant to  $\Psi$  to a large extent determines the nature of  $\Psi$  qua target of factual understanding. The constitutive role of purposes comes clearly to the fore in case, for instance, of understanding complicated vs. simple artefacts. We are prone to conceive of more complicated artefacts as agents, or agent-like entities, *by virtue of* our aim to explain and predict their behavior in terms of desires and intentions. Simple artefacts, by contrast, tend to be categorized as physical objects simply *because* we usually purport to describe them in physical or mechanistic terms.<sup>131</sup> Which purposes we choose to pursue (e.g. prediction and explanation in

---

<sup>130</sup> This is not to say that there is no epistemic merit to being aware of some piece of information's relevance to  $\Psi$ . I would account for such awareness (or rather, for its behavioral implications) in terms of having factual understanding not of  $\Psi$ , but of the relation between  $\Psi$  and the information at hand.

<sup>131</sup> This is, of course, reminiscent of Dennett's view of intentionality (Dennett, 1987). But it is also more general than Dennett's view. For while Dennett was concerned only with

intentional language vs. description in mechanistic terms) is, at base, just that: a choice. But it is not a random choice: the pursuit of the adopted purposes must contribute sufficiently towards the overall goal of factual understanding: optimizing a community's attunement to its environment. So, in sum, the externalist reading of the two provisos ultimately boils down to a pragmatist account of what it means for something to concern, or pertain to,  $\Psi$ . The facts that determine what information, or which purposes, are relevant to  $\Psi$  are facts about what a community at large considers relevant to  $\Psi$ . And what a community considers so relevant is determined, in turn, by the community's overall purpose to remain optimally attuned to its environment.

### 2. *The relation between sender and recipient*

In order to have factual understanding, I have maintained that an agent – the sender – should be disposed to transmit information to another agent – the recipient. This invites the question to what extent informational transmission requires the sender to be acquainted with the recipient. Does the definition cover only paradigmatic, ‘eye-to-eye’ scenarios, or does it also apply to cases in which the sender is unaware of the existence, let alone the identity, of the recipient? I opt for a very tolerant interpretation of the acquaintance requirement, to the effect that  $A$  need not be directly acquainted with  $B$ , but must at least tacitly entertain certain beliefs or assumptions about  $B$ . These beliefs or assumptions must be such that it makes sense for  $A$ , first of all, to take himself to be transmitting information to  $B$ , and secondly, to impute certain purposes to  $B$ . This means, for instance, that a lecturer can be said to possess factual understanding by virtue of being disposed to transmit information through online means to a muted and invisible audience. In order for this understanding to be had, the lecturer must at least act on the supposition that there *is* an audience, and must have a general idea about the nature of the audience (e.g. that it consists of human beings who are familiar with the language of instruction), so that it makes sense, from the lecturer's perspective, to impute certain purposes to the audience to which the transmitted information is relevant.

The tolerant interpretation of the acquaintance requirement can be extended to cases in which not only the nature, but also the very existence of the recipient is at issue. This accommodates the possibility that agents may possess factual understanding by virtue of writing articles and books to presumed (but by no means necessarily existent) readers. It can even accommodate the idea that factual understanding manifested in the information carried by the Voyager spacecraft, which we have some reason to believe will *never* reach its hoped-for recipient(s). All this permissiveness is grounded not upon intuition, but upon the consideration

---

marking a distinction between agents and non-agents in terms of the types of descriptions we deem relevant to understanding them, I am suggesting a view according to which our descriptive purposes are formative of the way in which we categorize *all* kinds of targets of factual understanding.

#### 4 Factual understanding

that in order to most effectively realize the general objective of the practice of factual understanding, agents within a community should be maximally incentivized for sharing environmental information. That is, a community should want to reinforce information-sharing behavior within that community as much as possible, through letting the domain of understanding-attribution be as large as possible given the constraints the overall goal of the practice of factual understanding imposes on that domain. This means that as long as there is at least a *potential* purpose to which the sender's information could serve as a means for a *potential* recipient, then the sender should in principle be liable to an ascription of (some degree of) factual understanding. In part, this desideratum was already heeded by countenancing possible (in addition to actual) purposes of recipients (see condition (ii): "has *or* could have"). The tolerant interpretation of the acquaintance requirement, which countenances actual, presumed, as well as merely desired recipients, constitutes an additional route via which the desideratum is taken account of.

#### 3. *Putting agents in a position to achieve a purpose*

The third and last target of clarification concerns a phrase featuring in clause (b) of condition (ii). As (iib) says, the sender should aim to put the recipient in a position to achieve some purpose pertaining to  $\mathcal{P}$  which the sender attributes to her. As clarified earlier, condition (iii) then demands that this purpose – the enabling of the recipient – should be achieved on the basis of the transmitted information. There are at least two questions that may be raised with respect to the notion of enabling. First, is there a universal 'threshold' which demarcates non-enabling from enabling transmissions? And second, to what extent does enabling hinge on the (cognitive) abilities of the recipient? I will attempt to answer these questions in turn. To be clear, my answers will not yield a comprehensive account of what the phrase 'putting one in a position to achieve some purpose' means. But I do hope to at least eliminate any problematic vagueness lingering in the unclarified version of that phrase.

As to the first question, I would argue against the existence of a universal threshold, insisting instead on the inherent graduality of the notion of enabling. In fact, this inherent graduality coincides with the fourth parameter identified at the end of the previous section: specificity. A sender can enable the recipient to do certain things to different degrees of effectiveness, in the sense that a sender's transmission can be strongly or weakly tailored to the achievement of a purpose attributed to the recipient. Perhaps the purpose for which the sender intends the recipient to use the transmitted information, requires the recipient to engage in a very demanding sequence of reasoning – a scenario in which the sender's transmission would be weakly tailored to the attributed purpose. Or in the opposite scenario, the transmission might be so specific that the recipient only has to comply with detailed instructions contained in the transmission in order to achieve the intended purpose. These examples are merely illustrative, however. What exactly determines how well a message is tailored to some purpose is a matter I leave open here; I merely wish to

point out that insofar as the vagueness of the notion of enabling is due to the inherent graduality of that notion, my account (i.e. the fleshed-out version of it) is designed to accommodate that graduality in terms of the parameter of specificity.

My answer to the second question is best introduced by means of a simple example. Suppose that someone (agent *A*) is disposed to explain, in Mandarin, Newtonian mechanics to a monolingual speaker of English (agent *B*), whose total lack of understanding of Mandarin is not apparent to *A*. Can *A*, in this case, be said to enable *B* to achieve any purposes relating to Newtonian mechanics? Arguably not. Then again, *A*'s explanations need not be altogether void of epistemic merit. The explanations might enable other possible recipients (e.g. speakers of Mandarin) to achieve certain purposes relating to Newtonian mechanics. Were *A* to have directed his transmission at a more suitable recipient, he would have been liable to an ascription of factual understanding. It would seem overly harsh to deny *A* factual understanding solely because his explanatory efforts are misdirected. Not so much because of any intuitive judgment to that effect, but rather because *A*'s information may potentially be very useful to others – just not to the actual agent who happens to be the recipient of that information. In light of the main purpose of ascribing factual understanding, there should in principle be room for crediting *A* with understanding Newtonian mechanics in the sketched scenario.

Fortunately, the permissive conception of the relation between sender and recipient (see clarification #2) allows us to reconstruct cases like the above in a way that yields a more charitable assessment of *A*'s epistemic status. The reconstruction is the following. Rather than at *B*, *A* is implicitly directing his transmission at a different recipient, namely at a counterfactual agent *B\** who does comprehend Mandarin but is otherwise identical to *B*. Relative to *B\**, *A* can be said to have factual understanding of Newtonian mechanics, provided of course that *A*'s transmission does indeed put *B\** in a position to achieve the purpose attributed to her. *A*'s misjudgment of *B*'s linguistic capabilities can then be construed as an epistemic deficit concerning a different form of factual understanding, namely the factual understanding *A* has of *B*. For it is with regard to his understanding of *B* that *A* would fail to put others in a position to achieve certain purposes regarding *B*. For instance, were *A* to aim to transmit information concerning *B* by uttering the sentence '*B* understands Mandarin', recipients of that transmission would not be able to achieve the purpose of, let's say, successfully predicting how *B* would respond to being addressed in Mandarin.

Other examples of 'misdirected transmission' can be reconstructed in similar fashion. Consider a situation in which a philosophy professor offers a highly abstract exposition of Kant's ethics, thereby puzzling rather than successfully informing his students. Based on my reconstruction of the previous example, the professor's factual understanding of Kant's ethics can be salvaged if there are possible (philosophically more well-versed) recipients to whom the professor's transmission would be useful in the intended way, and to whom, by that token, we should take the

#### 4 Factual understanding

professor to be implicitly directing his transmission. The professor's failure to successfully educate his actual students can then be seen as rooted in a lack of factual understanding of his students' abilities and background knowledge.

My reconstruction procedure offers a way of understanding the notion of enabling in such a way that ascriptions of factual understanding do not depend on properties and abilities of recipients to the extent that we would need a detailed account of enabling in order to make proper sense of the implications of my definition. The reconstruction procedure hinges on the following principle of charity: when assessing some agent's level of factual understanding, identify the recipient in a sender-favorable way, allowing in some cases the ostensible, actual-world recipient to be substituted by a possible recipient. The upper bound of legitimate substitution is constituted by the requirement (cited under condition (iii)) that the achievement of the sender's purpose be due to his transmission. In other words, the sender's transmission should figure ineliminably in an explanation of how the recipient is put in a position to achieve the purpose attributed to her. So a scenario in which the recipient (actual or possible) first has to negate or drastically modify the content of the sender's transmission does not suffice for an ascription of factual understanding. In such a case, after all, the transmission does not figure ineliminably in an explanation of how the recipient is enabled to achieve a certain purpose. But apart from this upper bound, the procedure is quite flexible – a flexibility which is rooted upon a kind of 'every-little-helps-principle': whenever someone's informational transmission regarding some  $\Psi$  is potentially useful to a community, a community should want to foster the sustainment of that disposition.<sup>132</sup> Any deficiencies of that disposition (e.g. its misdirectedness) can be taken care of by denying agents factual understanding of their actual recipients, thus discouraging the formation (or encouraging the adjustment) of dispositions for transmitting information about those recipients.

#### 4.4 Replies to objections

In this section I discuss four objections that may be levelled at my proposed account of factual understanding. The first two objections constitute challenges to what may be called the internal plausibility of that account. According to these objections, my proposal does not succeed in achieving its own aims, even if we accept its main assumptions. The final two objections cast doubt on the external plausibility of my

---

<sup>132</sup> Of course, there are limits to the extent to which a community can anticipate the potential usefulness of information. It could be that a transmission of information which is completely incomprehensible to all current members of a community, turns out to be highly useful to future members of that community. In such a case, in which the community's current members are unable to foresee the information's future usefulness, the sender should not be ascribed factual understanding purely out of a hopeful expectation that his information could someday find its proper audience. This is because such a lenient ascription protocol would render the practice of ascribing factual understanding insufficiently effective in discouraging the dissemination of information which is, and will continue to be, useless.

account; challenging the idea that that account could live up to the requirements anyone may reasonably impose upon a theory of factual understanding. My aim is to defeat each of the four objections.

*1. Criterion (ii) is redundant. A purpose to transmit information regarding worldly targets is already implicated in the competent use of a symbol (i.e. criterion (i)).*

My account of factual understanding is built on the assumption that the use of a symbol does not in and of itself entail transmission of information. The above objection, however, says that when one is disposed, through learning, to use a symbol properly (i.e. for its associated purpose), then one is *ipso facto* disposed to transmit information about the world. I presume that one could support this claim by arguing that the significatory purpose of symbols, or at least a certain class of symbols, is inherently transmittory: it is in the nature of symbols to convey information about the world. Take the sentence ‘Amsterdam is the capital of the Netherlands’. Is it not the case that the proper use of this sentence would automatically entail making an assertion *about* the capital of the Netherlands? And would this not constitute transmission of information?

My counterargument to this challenge is best introduced by means of a rather pedestrian example. Consider an everyday household item, such as a cheese slicer. I think we can agree that a cheese slicer is an object with a determinate function, and that a cheese slicer is therefore the kind of thing that can be used properly (e.g. in accordance with its function). Suppose that one has a disposition to use a particular cheese slicer properly.<sup>133</sup> This, I take it, amounts to being disposed to use the cheese slicer for the *immediate purpose* it exists for: the purpose to facilitate the production of slices of cheese of a certain thickness. Now, it seems clear that slicing cheese is intimately connected also to the following, *associated purpose*: the preparation of cheese for further use – typically consumption. It is at least difficult to imagine a reasonably plausible scenario in which one uses a cheese slicer properly (i.e. in order to slice cheese), but not for a purpose to prepare the cheese for further use. One might conclude, therefore, that to slice cheese *is* to prepare the cheese for further use. On this view, the immediate and the associated purposes of a cheese slicer are two sides of the same coin.

I resist this view, however. For what appears to be an unassailable connection is but a constant conjunction of purposes. Even if the two purposes invariably occur together, there is still a sense in which they can be kept apart. Indeed, there is a value in keeping them apart: discriminating between the two purposes allows us to explain, in a means-to-end sense, the one in terms of the other. It is the associated purpose which explains the immediate purpose: the fact that we have a need for being assisted

---

<sup>133</sup> As one may know from experience, having a disposition for proper usage of one particular cheese slicer does not always imply having such a disposition with regard to other cheese slicers!



#### 4 Factual understanding

in preparing cheese for further use explains why we have created for ourselves items which facilitate the production of slices of cheese. The achievement of the cheese slicer's immediate purpose is a means to achieving its associated purpose. This accounts for why the use of a cheese slicer is (barring highly unusual examples) *invariably* associated with the preparation of cheese for further use, but not *inviolably* so. The connection only appears to be inviolable because cheese slicers have been designed in order to enable us to achieve their associated purpose, and are therefore, more or less without exception, used in connection to that purpose. Ultimately, however, a cheese slicer's associated purpose is not strictly constitutive of its immediate purpose. The relation between the cheese slicer's immediate and its associated purpose is still of a means-end type, and hence contingent. In principle, we could conjure up a different associated purpose – one in which slices of cheese do not figure as objects of further use – and choose to utilize cheese slicers for that purpose whilst continuing to attribute it with its original immediate purpose.

A somewhat similar story can be told with regard to the use of symbols. As I have argued in section 3.1, the practice of ascribing symbolic understanding exists in order to ensure that a community can maintain communicative efficiency. Ascriptions of symbolic understanding serve to reinforce certain behavior, and discourage other behavior, in such a way that there is continuity in a community's system of symbols and interpretative rules. This conception of the goal of the practice of symbolic understanding has implications for how we should think of the purpose – the *immediate* purpose – of the use of symbols. I have said in section 3.2 that the purpose of a symbol consists in the type of outcome to which compliance with its associated rule of interpretation tends to give rise. For convenience, I recite the example that featured in that section (p. 88):

For instance, if you say 'I am Chris', expected consequences of producing this utterance are that others subsequently address you as 'Chris', use that name to refer to you when talking to mutual acquaintances, and expect you to reliably respond in certain ways upon being called 'Chris'. Such are the consequences of introducing yourself.

I have also explained that the proper use of a symbol amounts to selecting it for its purpose. Thus, proper usage of the sentence 'I am Chris' would consist in selecting it (e.g. uttering it) in order to realize the effects cited above. But it is here that a further subtlety can be discerned. Given that ascriptions of symbolic understanding serve to reward the proper usage of symbols, and to discourage their improper use, such ascriptions effectively stimulate agents to realize certain outcomes *insofar* as those outcomes corroborate their having complied with the appropriate interpretative rules. An agent who is disposed to use a symbol properly thus, in effect, has a disposition to seek (implicit) acknowledgement of the viability of her own interpretative habits.

The point I am trying to make is that symbolic understanding, in contrast to factual understanding, consists in self-directed rather than in other-directed behavior. An agent's immediate purpose in using a symbol properly is to have her own interpretations confirmed by the behavior of other agents. It is as if the utterer of 'Amsterdam is the capital of the Netherlands' issues a request at her recipient of the form 'Please now act in a way which confirms my expectations of what effects the use of 'Amsterdam is the capital of the Netherlands' should have'. Irrespective of whether the recipient is able to achieve any purposes *he* might have in interpreting the utterance (such as being able to pinpoint the capital of the Netherlands on a map), the immediate purpose of the utterer is accomplished as long as the recipient starts acting in a way that matches the expectations of the utterer. To be sure, such confirmation of the utterer's expectations may come in the form of the recipient's accomplishing his goals. But it is not towards the accomplishing of those goals that the utterer's speech act is directly geared. The accomplishing of the recipient's goals is collateral to the utterer's immediate purpose: acquiring confirmation of her expectations.

Where the symbol-user's *immediate* purpose is thus self-directed in the sense that she aims to have her expectations confirmed, her *associated* purpose may, and usually will, be other-directed. Just as we tend to slice cheese in order to prepare it for further use, we tend to use symbols in order to have those symbols used for further purposes. When such a further purpose concerns the achievement of goals having to do with worldly targets (*Ψ*s), we enter the domain of factual understanding. As with cheese slicers, the connection between the immediate and the associated purpose of using symbols is strong but not unassailable. It is strong because the associated purpose explains the immediate purpose, but not unassailable because the explanatory relation is of a means-to-end kind, and hence ultimately contingent. Just as our need for preparing cheese for further use explains why we have cheese slicers, our need for having a practice of factual understanding explains (at least in part) why we have a practice of symbolic understanding. That is, the need to promote the dissemination of information within a community requires that there is some stable system of symbols and interpretative rules in terms of which information can be transmitted. The existence of such a system allows agents to issue the highly implicit requests that instances of symbol-usage are, and to have those requests complied with. All this is meant to show that the objection does not succeed: while a purpose to transmit information regarding worldly targets is strongly associated with the competent use of symbols, it is not entailed by it. Consequently, condition (i) is not redundant.

#### 4 Factual understanding

*2. The definition is too restrictive: in order to have factual understanding, one need not have a purpose to transmit information (and thereby let others achieve certain goals).*

This objection amounts to another challenge from within. It arises from the consideration of scenarios in which the successful performance of some action by an agent enables others to achieve certain purposes, without this being the intended effect of the agent's performance. For instance, it seems one can unintentionally convey understanding of a complicated machine to others through operating it successfully, without being aware of the presence of onlookers – let alone having a purpose to inform them. Should we not deem the operator to possess factual understanding of the machine, or of some operating procedure, by virtue of this inadvertent transmission of information? My answer is no, a purpose to transmit information is necessary for factual understanding. The reason for this is that without the requirement of purposefulness, attaining the main goal of the practice of factual understanding would become significantly less feasible.

I acknowledge that the above objection has a clear intuitive appeal: provided that one buys into the general behaviorist framework underlying my account of factual understanding, the criterion that transmissions be purposive may indeed seem intuitively redundant. However, intuitive plausibility is not the primary standard of theory evaluation here. Instead, my account should be evaluated in terms of how well it accounts for what I have identified as the main goal of the practice of ascribing factual understanding: optimizing the attunement of a community (and hence agents and groups within the community) to its environment. Relative to this evaluative standard, purposiveness of informational transmission is crucial. It is crucial, because without such a criterion the practice of ascribing factual understanding would be much less effective. That is, a practice in which the notion of factual understanding would apply to purposive and non-purposive transmissions alike, would promote not information-transmitting behavior as such, but rather a much larger category of behavior from which other agents may or may not be able to reliably extract useful information.

Let me elaborate on this. I have said in chapter 1 that ascriptions of understanding (symbolic and factual) are to be seen as means to encouraging the possession of certain dispositions. On this picture, to say that 'A understands  $\Psi$ ' is to say something like 'Excellent, A! Please continue having disposition  $x$  in the future', or 'Hey folks, do what A is doing'. In case of factual understanding, the aim is to encourage agents to optimally contribute to the achievement of others' purposes. Now, if ascriptions of factual understanding were insensitive to whether a sender was purposefully or inadvertently transmitting information, those ascriptions would, as it were, underspecify the desired disposition. For example, if the machine operator were credited with factual understanding regardless of whether her performance was other-directed, this would not allow us to evaluatively discriminate between behavior that is appropriately tailored to potential purposes of others, and behavior

that is not so tailored. An ascription of factual understanding of this insensitive sort would incentivize the machine operator to continue operating the machine as she did before, but it would *not* incentivize her to calibrate her operations to the purposes of others in relevant contexts (namely in those contexts where others might, now or in the future, benefit from the operator's performances). Without providing a means to incentivize agents to attune their behavior to the possible purposes of others, the practice of ascribing factual understanding would not be as effective as it could be. This is because it would leave agents largely to their own devices regarding the way in which they had best interpret and utilize the observed behavior of others for their own interests. Unlike a community in which the ascription of factual understanding *were* sensitive in the indicated way, members of a community with an insensitive conception of factual understanding would not be able to benefit from the kind of incentivizing of other-directed behavior that the sensitive conception of factual understanding provides for.

There is a further way in which purpose-insensitive ascriptions of factual understanding would underspecify the desired effect of applying that concept. This has to do with the fact that the extent to which an agent's skills and abilities are internally anchored in learning, matters to how well others can benefit from the manifestation of those skills and abilities. What I mean by this is best illustrated by means of our machine operator. Suppose that a community awarded understanding to the operator solely on the basis of her disposition to successfully operate the machine. In this scenario, it could be the case that the operator's disposition, although superficially commendable, was rooted in what may be called 'irretrievable learning'. That is, the operator might have an efficacious disposition to pull a certain lever in order to activate some part of the machine, but this disposition could be rooted in mere muscle memory: the meagre behavioral residue of a disposition once acquired via some learning process, the contents of which now utterly elude the operator. If others were to have to benefit from observing the doings of the operator, a lot would depend on *their* ability to correctly identify which aspects of the operator's actions were responsible for which effects. For given the operator's 'acquired cluelessness', she would not, or at least not consistently, be able to attune her performances to the goals of others, even if she were properly motivated.

Without a requirement for purposeful transmission, a community would have insufficient means to eliminate the kind of acquired cluelessness that is due to irretrievable learning. It would have to make do with a concept of understanding the use of which promoted behavior from which others might not always be able to optimally benefit. The requirement that agents purposefully transmit information, however, sees to it that agents, so to say, 'know what they are doing'. For in order to understand the machine on such a sensitive notion of factual understanding, the operator must not only be able to operate it successfully, but she must be able to isolate the sequence of behavior pertinent to the successful operation of the machine. After all, if the operator's aim is to enable others to achieve certain purposes relating

#### 4 Factual understanding

to the machine, she must be able to identify – at least implicitly – which aspects of her behavior are responsible for which effects.

It should be noted that the requirement that understanders ‘know what they are doing’ does not amount to what is commonly called a *transparency condition* on understanding. According to this condition, to understand something requires being aware, or having understanding, of one’s understanding. Such a strong transparency requirement is not what I have in mind with the desideratum that agents ‘know what they are doing’. The quotation marks are important: they indicate that the phrase is not to be taken literally. Within the context of my account of factual understanding, knowing what you are doing requires that you are able to identify the ‘active ingredients’ that are instrumental to achieving the purpose you attribute to another agent in transmitting information. Such identification does not demand awareness, let alone understanding, of the ingredients identified. It does not demand this because of my externalist account of purposefulness (section 2.1). On that view, acting for a purpose does not necessitate the presence of a mental state such as an intention, or a desire-belief pair. Instead, it only necessitates the presence of a disposition to have one’s behavior converge to a particular type of outcome. Rather than requiring conscious awareness, ‘knowledge of what one is doing’ is therefore simply implicated in the disposition for informational transmission required for factual understanding. More concretely, that ‘knowledge’ is implicated in the sender’s purpose to have her recipient interpret her transmission in a certain way. For example, if a machine operator purports to enable others to activate a certain part of the machine through pulling a lever, this requires her to purposefully signify (=symbolize) the relevant sequence of behavior (the pulling of the lever) in an act of informational transmission.<sup>134</sup> At most, the knowledge of what one is doing involved in such purposeful signification amounts to a very weak form of transparency. See also the discussion under the next objection.

*3. The definition is too permissive: in order to have factual understanding, one needs to be aware, or indeed understand, that one has factual understanding.*

This objection takes up on the final issue discussed above. Rather than constituting a challenge from within, this objection starts from a pretheoretic, intuitive conception of understanding. As such, it can be restated in terms of the complaint that my explicatum of factual understanding strays too far from its associated explicandum, and therefore violates Carnap’s similarity criterion. That explicandum, according to those sympathetic with the complaint, harbors a transparency criterion which says that understanding (unlike knowledge) must come with higher-order understanding: to understand is (also) to understand that one understands. As Zagzebski (2001, p.

---

<sup>134</sup> This can be done verbally (through describing the procedure), as well as non-verbally (through demonstrating how the lever is to be pulled). Both involve symbolization. See also my comments about the driving instructor in section 4.2.

246) puts it: “[U]nderstanding is a state in which I am directly aware of the object of my understanding, and conscious transparency is a criterion for understanding.” In somewhat more nuanced terms, Pritchard (2010, p. 76) writes: “Understanding clearly is very amenable to an account along epistemically internalist lines, in the sense that it is hard to make sense of how an agent could possess understanding and yet lack good reflectively accessible grounds in support of that understanding.” Apart from the purported intuitive plausibility of this idea, the transparency condition on understanding is sometimes thought to procure a welcome demarcation from the concept of knowledge, and to help account for understanding’s distinctive epistemic value (Grimm 2017).

I have the following to say in response to the challenge that my account of factual understanding is defective because it lacks a much-needed transparency requirement. First of all, I wish to repeat that observing fidelity to our intuitions is *not* the main methodological principle of this investigation. In order to assess the viability of including a transparency requirement in the explicatum, the question should rather be: does the addition of such a requirement make better sense of how a practice of ascribing factual understanding contributes to the overall goal of optimizing a community’s attunement to its environment? The answer, to my mind, is no: an all-out transparency condition (along Zagzebski’s lines, that is) would be redundant. As explained earlier, transmitting information with an eye to enabling others achieving certain purposes, demands that there should indeed be a rudimentary form of transparency, in the sense that the agent can only tailor her transmissions to the purposes of others provided that she can pinpoint that which is instrumental to achieving those purposes. Given that the notion of a purpose is externalistically defined (i.e. as a type of outcome to which behavior converges), such pinpointing need not consist in a mentally accessible state or process. Instead, it can simply be implicated in engaging in the purposeful act of informing others. Transparency, in the strong sense, is thus not required for factual understanding. In fact, inclusion of a transparency condition would to some extent thwart the goals of the practice of ascribing factual understanding. It would give rise to an overly exclusionary practice of ascription, in which perfectly reliable informants would sometimes fail to be justly reinforced in their information-transmitting behavior. Hence, there is an argument to be made for the case that a transparency requirement would not be just redundant, but also too restrictive.

But, one might counter, is it not the case that having reflective awareness of one’s understanding is a commendable quality that warrants some kind of epistemic appraisal? I agree, and I think this can be accounted for from within the purview of my account of factual understanding. My take on the matter is that transparency (that is, in the strong, Zagzebskian sense) is epistemically commendable insofar as it gives rise to behavioral dispositions relating to the transmission of information about one’s own epistemic status. If an agent is disposed to usefully inform others about her own

#### 4 Factual understanding

understanding of some target  $\Psi$ , then she may be credited with having factual understanding of *her understanding of  $\Psi$* .

Let me illustrate this. Consider, again, our machine operator. Suppose now that she has strong reflective awareness of how her manipulations of the machine conspire to bring about certain effects. I have argued that such awareness does not impact on her understanding of the machine itself. For with respect to *that* understanding, the operator's reflective awareness is supererogatory. But her awareness *can* impact on the understanding the operator has of her own understanding of the machine. It does so in case the operator's reflective awareness gives rise to a behavioral disposition for enabling others to achieve certain purposes with regard to the operator's understanding of the machine. For instance, the operator may have a purpose to let others successfully anticipate her machine-related behavior, as a concomitant purpose of her effort to explain how certain manipulations of the machine lead to certain outcomes. This is different from her simply conveying understanding of the machine itself. In case the operator is merely disposed to explain the workings of the machine for the purpose of letting others achieve machine-related goals, she only has factual understanding of the machine. But in case the operator has a (concomitant) disposition to successfully inform others about her mastery of the machine, this amounts to factual understanding of her own understanding of the machine. Transparency, provided that it procures behavioral dispositions of the right kind, is thus instrumental to a special, 'meta'-variety of factual understanding.<sup>135</sup>

*4. The proposed account of factual understanding is implausible because it does not include a veridicality requirement.*

This objection, which is arguably the most damning criticism that may be levelled at my proposal, can be specified in the following way. Since the definition of factual understanding makes no reference to truth-as-correspondence, but only to the achievement of purposes, the account is not equipped to deal effectively with an important class of examples. For instance, the account seems to entail that one can have understanding through the transmission of incorrect but nonetheless useful information. Worse still, it also appears to imply that the transmission of incorrect, but useful information gives rise to a higher degree of understanding than the

---

<sup>135</sup> So, to be clear, it is not transparency as such which is an object of epistemic appraisal, but rather the behavior it renders possible. This ties in with what was said about the role of cognition for symbolic understanding in section 3.2. Within my framework, cognition is of undeniable epistemological importance, but only qua means to the formation and sustainment of behavioral dispositions. The target of an ascription of symbolic or factual understanding is always a behavioral disposition, because the general goal of ascribing symbolic and factual understanding is to reinforce and curtail behavioral dispositions. It is only in a derivative sense, therefore, that cognitive states and dispositions (including those involved in transparency) are units of epistemic assessment.

transmission of correct, but less useful information. Surely such implications are to be avoided? And surely that is to be done by imposing a correspondence-based truth condition upon informational transmissions? It would seem, after all, that only such a condition can effectively cancel the counterintuitive implications just mentioned. Below, I respond to this objection, first, by pointing out that the addition of a correspondence-based truth requirement would not sit well with the general methodological principle on which this investigation is predicated. Then, I argue that my account is in fact equipped to deal with the aforementioned types of examples. I close off by discussing an important consequence of my rebuttal of this objection.

Intuitively, one's understanding should align with a mind-independent reality, or at least approximately so. The inclusion of a correspondence-based truth criterion would answer to this demand. It would allow us to pass eminently reasonable verdicts, such as that combustion cannot be understood in terms of phlogiston theory, because phlogiston theory is false. It would also enable us to capture the intuition that usefulness is no varnish for falsehood; even if phlogiston theory were useful in some way, combustion could still not be understood in terms of it. I contend, however, that a correspondence-based notion of truth, while widely adhered to in the contemporary literature on understanding, should not figure as a condition on factual understanding. As my qualms with that notion are similar to those voiced by Peirce, it is worthwhile to briefly dwell upon Peirce's concerns. As Misak (1991, p. 39) explains, Peirce considered the correspondence theory of truth to be "pragmatically spurious", because it relies on a transcendental idea of reality as an "unknowable thing-in-itself":

Peirce suggests that the notion of an unknowable thing-in-itself has no consequences for ordinary experience. We can say nothing about it: 'The Ding an sich . . . can neither be indicated nor found. Consequently, no proposition can refer to it, and nothing true or false can be predicated of it. Therefore, all references to it must be thrown out as meaningless surplusage.' (CP 5. 525, 1905). The correspondence theorist cannot say that the world as it is in itself has properties *a* and *b*. So if truth is defined as correspondence with that world, no expectations can be derived from '*H* is true'. If we do not know what correspondence with the world would be like, we cannot know what to expect of hypotheses which so correspond. A theory of truth that centers around something unknowable utilizes 'ghost-like hypotheses about things-in-themselves which anybody can set up but nobody can refute' (CP 7. 370, 1902). Since an account of truth purports to be about the world, it must have consequences for experience—not experiences in special experimental setups, but experiences which we can all have. The correspondence theory does not have such consequences and thus it is spurious.

Although I do not explicitly side with Peirce regarding his own, positive conception of truth as that "which would be believed at the end of inquiry" (Misak, 1991, p. xiv), I do think that his critique of the correspondence theory is of relevance to the



#### 4 Factual understanding

present project. The relevance resides in the fact that I make a similar attempt at providing a “pragmatic elucidation” (idem, p. viii) with regard to understanding as Peirce was trying to do with respect to truth. In such a project, there is no room for concepts which lack practical consequences.

Some further elaboration on this point is due. My aim in this chapter has been to explicate factual understanding in such a way that the explicatum makes sense of the general purpose that ascriptions of factual understanding serve. That general purpose, I have assumed, consists in optimizing the attunement of a community to its environment. I understand such attunement in a practical sense: a community is more attuned to its environment the more the community at large, and individuals and groups within the community, are able to achieve their goals through engaging in efforts to achieve those goals. One could object that this practical notion of attunement is ill-conceived, as it fails to exclude cases we intuitively want to exclude (e.g. having understanding based on phlogiston theory). For this reason, it would seem advisable to replace the practical notion with an alethic one, according to which attunement correlates with truth: the more agents within a community entertain true beliefs, or the more they utter true statements, the more they are attuned to their environment. But it is here that the Peircean concerns enter: if this truth-based form of attunement were the general purpose of ascribing factual understanding, then it would be unclear under which circumstances factual understanding could be ascribed. Attunement-as-truth is not a type of outcome the occurrence or non-occurrence of which impacts on our experience; it is a spurious, or empty purpose. Such a notion of attunement can therefore not figure in an account that attempts to make sense of what the concept of factual understanding is for.<sup>136</sup>

One might still wonder, however, how we should deal with the counterexamples my truth-eschewing account of factual understanding faces. Clearly, it would not do to argue that theories commonly regarded as false are also without practical merit. For many ostensibly false theories do enable us to achieve certain purposes. In fact, some false theories may be, or may once have been, more useful than true(r) ones. As Wilkenfeld (2017, p. 1278) reminds us: “[W]hen Copernicus first proposed his heliocentric model, its predictions were actually less accurate than the Ptolemaic alternative.” Notwithstanding my earlier insistence that intuitive plausibility is not the appropriate standard of evaluation here, the implication that the use Ptolemaic astronomy therefore would, at least back in the days, have given rise to a higher degree of factual understanding than the use of Copernican astronomy, is too

---

<sup>136</sup> One could argue that my practical notion of attunement amounts to a pragmatist conception of truth, and that therefore, the explicatum of factual understanding includes a veridicality requirement after all – albeit not a correspondence-based one. I am not opposed to this idea per se, but refrain from committing to it out of caution. For in order for me to justifiably claim that practical attunement = truth, I would, in keeping with my own methodological principles, have to (1) identify the function of truth-ascriptions, (2) explicate truth in accordance with that function, (3) and show that the resulting explicatum equates my notion of attunement.

uncomfortable to be ignored. Fortunately, this uncomfortable implication can be made more palpable by seeing that purposes are ever dynamic. The Copernican system may not have immediately performed better at, say, predicting the apparent retrograde motions of the planets, compared to the Ptolemaic system. But, at a later stage, it did turn out perform better with regard to predicting the phases of Mercury and Venus and the existence of stellar parallax – among other things. Those later discoveries, however, were predicated on the use of instruments not available at the time Copernicus’ findings were first published. In a sense, Galileo Galilei’s observations of the phases of Venus, and Friedrich Bessel’s confirmation of stellar parallax were achievements of novel purposes that could only be formed, and pursued, thanks to the development of the telescope and the heliometer. So relative to the purposes of a 16<sup>th</sup>-century community, the use of Ptolemaic astronomy afforded a higher degree of factual understanding. But relative to the purposes of later communities, the Copernican clearly fared better.

Admittedly, there is a kind of ‘diachronic relativism’ that results from this picture. It implies that theories that we do not currently consider to confer high degrees of understanding may once have done so, because the purposes of our predecessors differed from ours. The fact that present-day Ptolemaic astronomers do not qualify for ascriptions of factual understanding of the planetary motions, does not imply that their 15<sup>th</sup>-century colleagues did not qualify for such ascriptions in *their* time. I see no harm in such relativism. I grant that at an intuitive level, we might want ascriptions of understanding to be diachronically fixed. But such a ‘timeless’ conception of understanding would be tangential to the idea that the application of a concept should have practical consequences. If we assume, as I have done, that the concept of factual understanding is to be explicated in such a way that the explicatum makes sense of the general purpose the concept serves for a community, then an atemporal, context-insensitive notion of understanding is neither here nor there. Information-transmitting behavior can only ever be reinforced or curtailed against the background of a set of purposes that a community has at a certain point in time. To demand that ascriptions of factual understanding be diachronically fixed is therefore like demanding that ascriptions of factual understanding track the truth: it is a spurious purpose. It is spurious, since we have no means of ascertaining under what circumstances it is or is not achieved.

#### **4.5 Benefits of the explicatum**

Explications are justified in a means-to-end sense: they are vindicated by the goals that can be accomplished by means of them. As Carnap wisely emphasized, however, not all ends justify all means: any violations of the similarity criterion must be shown to be sufficiently inconsequential or otherwise unobjectionable. Furthermore, those violations must be necessary in the sense that the explication’s aims would not be attainable without them. So far, I have been at pains to show that my explicatum of factual understanding diverges from its associated explicandum in ways that are

#### 4 Factual understanding

defensible. In this section, I will highlight some of the advantages of the explicatum. In particular, I hope to show how my account of factual understanding can cast light on some of the issues that initiated the ‘turn to understanding’ in epistemology, as well as on other subjects that have featured prominently in the understanding literature. The following five topics will be touched upon:

- 1) The relation between symbolic and factual understanding
- 2) The epistemic acceptability of idealization
- 3) The relation between objectual and explanatory understanding
- 4) Knowledge, understanding, and scientific progress
- 5) The epistemic value problem

The benefits that I will discuss, especially regarding the first topic, not only bear upon the legitimacy of the explicatum of factual understanding, but also upon the legitimacy of the explicatum of symbolic understanding. For it is because of the central role symbolic understanding plays in my account factual understanding, that any benefit accruing to the latter indirectly accrues to the former.

##### *1. The relation between symbolic and factual understanding*

There is a consensus in the epistemological literature that having symbolic understanding is in many cases crucial to possessing factual understanding.<sup>137</sup> In Baumberger’s (2014 p. 70) words: “Symbolic understanding is a precondition for factual understanding whenever the latter is symbolically conveyed.” And as De Regt & Gijssbers (2017, p. 50) put it: “To understand a phenomenon, (...) we typically need a representational device of the right kind which we then have to connect to the phenomenon in the right way.” While authors differ on how the relation between factual and symbolic understanding is to be understood exactly, the general idea that the latter is often or even typically essential to the former is uncontroversial.<sup>138</sup> Straightforward as this sounds, however, those who seek to

---

<sup>137</sup> Epistemologists tend to have their own ways of carving up the landscape of understanding. This means that one will rarely encounter explicit uses of the terms ‘factual’ and ‘symbolic’. This does not contradict my observation, however. For in disambiguating the term ‘understanding’, most authors do in fact allude to the distinction between factual and symbolic understanding in some way or other. It is, for instance, not uncommon for authors to give only a passing mention to the kind of understanding that pertains to linguistic items and symbolic representations, and to subsequently equate the category of objectual understanding with my category of factual understanding, deeming the latter to exhaust everything of epistemological interest contained within the grammatical category picked out by ‘to understand + a noun (phrase)’.

<sup>138</sup> One interesting and original proposal concerning the relation between factual and symbolic understanding is due to Elgin (2017). She argues that the link between the symbolic and the factual is captured by the notion of *exemplification*, which she understands as the referential relation that samples bear to what they are samples of. We understand the world by virtue of understanding symbolic representations which exemplify features of that world.

salvage our intuitive notions of factual and symbolic understanding are confronted with a problem in accounting for the relation between these two categories of understanding.

Here is the problem. At the intuitive level, symbolic understanding has a cognitive as well as a behavioral dimension. Factual understanding, on the other hand, seems unequivocally cognitive. Indeed, that kind of understanding would appear to be a paradigmatically ‘internal’ affair. This mismatch precludes one from offering a satisfactory account of how factual understanding relates to symbolic understanding. An example helps to better appreciate the issue. Intuitively, understanding a phenomenon in terms of a model requires one to understand the model. And intuitively, also, one’s understanding of the phenomenon is better the better one’s understanding of the model: better symbolic understanding is at least a contributing factor to better factual understanding. On this picture, someone who understands the model behaviorally as well as cognitively, will, *ceteris paribus*, have a higher degree of understanding of the phenomenon than someone who only understands it cognitively. Yet if factual understanding is unidimensionally cognitive, as our intuitions dictate it to be, this should not matter. So the intuition-based approach at once predicts and denies that one’s level of symbolic understanding invariably has at least some impact on one’s level of symbol-mediated factual understanding.

My approach offers a way out of this conundrum: by explicating both symbolic and factual in behavioral terms, we resolve the mismatch and regain the ability to explain how, exactly, symbolic understanding relates to factual understanding. My proposal was that symbolic understanding is necessarily implicated in factual understanding: factual understanding comes in the form of dispositions in which the dispositions associated with symbolic understanding are ‘nested’. To have factual understanding is to be disposed to act for a purpose which encloses the purpose of proper symbol-usage. Since dispositions for disseminating information are delimited by dispositions for proper symbol-usage (e.g. limitations to an agent’s competent use of a symbol may curtail his ability to disseminate information via that symbol), one’s level of symbolic understanding contributes to one’s level of factual understanding to the extent that the breadth or depth of one’s disposition for proper usage of a symbol impacts positively on one’s disposition for transmitting information via the use of that symbol. My account thus says that (1) having some degree symbolic understanding is necessary for having some degree of factual understanding, and that (2) one’s degree of symbolic understanding enhances one’s degree of factual understanding to the extent that the dispositions involved in the former impinge upon the dispositions involved in the latter. The immediate advantage of this view over

---

In this context, it is important not to confuse my use of the adjective ‘factual’ (=applying to the non-symbolic) with the adjective ‘factive’ (=being true). Elgin can be interpreted as offering a non-factive account of factual understanding.

#### 4 Factual understanding

the intuition-based view is that it does not make inconsistent claims about symbolic understanding's relation to factual understanding.

A further advantage of my approach is that it enables us to account not only for the relation between symbolic and factual understanding, but also for the reason why this relation should exist in the first place. Why, indeed, would symbolic understanding be a vehicle for factual understanding? Based on my theory, we can answer this question by pointing to the fact that factual understanding revolves around the transmission of useful information, and to the fact informational transmission implicates symbolic understanding (in the sense in which I have explicated that concept here). Indeed, the need for a concept which promotes information-transmitting behavior entails the need for a concept through which agents can coordinate their behavior amongst one another. For to be able to effectively communicate one's aim to let someone else achieve some purpose, there must be a system of symbols, the use of which comes with certain expectations which should reliably correspond with the subsequent behavior of the receiving party. The ascription of symbolic understanding within a community sees to the conservation of that system through promoting its implicit mutual endorsement among agents.

As a final note, it is worth pointing out that my account also allows us to appreciate the fact that the relation between symbolic and factual understanding is, in a way, bidirectional. Although factual understanding as such is not a precondition for symbolic understanding, the practice of ascribing factual understanding does impact in an important way on the practice of ascribing symbolic understanding. Changes in the environment, as well as changes within agents, groups and communities, bring about changes in purposes. New needs breed innovations and adjustments of all kinds, which in turn breed new purposes, creating yet further innovations and adjustments. What counts as superior or inferior factual understanding is therefore continuously subject to incremental change. Our symbol systems must keep track with this unabating evolution. Novel terms must be coined to refer to useful items recently identified or invented. And interpretative rules must periodically be updated in order to recalibrate them with the changed expectations and responses resulting from the 'pressure' that developments in the realm of factual understanding exert upon inter-agent behavioral coordination. Symbolic understanding, although in an important sense prior to factual understanding, is therefore at the same time also subservient to our overarching aim to remain attuned to our environment.

#### *2. The epistemic acceptability of idealization*

In chapter 1, I mentioned Catherine Elgin as one of the early advocates of understanding-centered epistemology. In various publications, including her latest book *True Enough* (2017), she has pointed out that given the undeniably impressive track record of the sciences, epistemology would do well to accommodate our most well-respected scientific achievements. This demands that epistemologists find some

way of countenancing what Elgin calls ‘felicitous falsehoods’: idealizations which evidently stray from the truth in one way or other, but which, equally evidently, are of eminent importance to scientific inquiry. This call has been heeded by various other prominent voices in the understanding debate, such as Henk de Regt and Angela Potochnik. As these authors point out, a standard, truth-centered epistemology is not well equipped to make appropriate room for idealization. For if we assume that understanding must be factive, as truth-centered epistemology would have it, then we are compelled to choose between two evils. Either we would have to say that idealizations are only epistemically acceptable insofar as they approximate the truth, or we would have to come up with complicated reconstructions according to which idealizations are in fact very elliptically stated, hedged truths. Neither strategy is very attractive. The first strategy implies that idealizations are more acceptable the less they deviate from the truth. But given that less veridical idealizations tend to be at least as crucial to our scientific achievements as more veridical ones, this strategy does a bad job of really salvaging those achievements. The second strategy is problematic at best. For even if we assume that the reconstructions are possible, we would still struggle to accommodate the fact that in actual scientific practice, the value of idealizations usually does not hinge on their being amenable to reconstruction.

Potochnik (2020, p. 941) notes that “there is clear evidence that idealizations are not merely tolerated in the vehicles of our scientific understanding but play central, positive roles. If so, scientific understanding is at least sometimes achieved not in spite of but (in part) because of a sacrifice of truth or accuracy”.<sup>139</sup> According to Potochnik (2015, p. 75), scientific inquiry is best conceived of as having a diverse and dynamic set of aims, rather than as being solely directed at truth:

[T]here are a variety of aims of science, both epistemic and non-epistemic. Traditionally appreciated aims include (at least) accurate prediction, explanation, and representation. Other aims of science have recently received increasing attention. These include providing information to guide policymaking (...); action within a short timespan (...); and facilitating the public uptake of scientific knowledge. (...) Successful pursuit of one among the various aims of science generally inhibits success

---

<sup>139</sup> Potochnik is by no means the only one to have pointed this out. To provide just one extra example, here is a quote from Elliott & McKaughan (2014, p. 6):

It is a descriptive fact of ordinary scientific practice that models represent their targets with varying degrees of success and typically focus selectively on those factors that are necessary in order to achieve the purposes for which they are used. Even our most successful models of nature are often known to be partial, simplified, incomplete, or only approximate. Indeed, the use of idealization, which goes beyond mere abstraction by deliberately employing assumptions known not to be true of the system of interest (e.g., treating bodies as point masses, surfaces as frictionless planes, collisions as perfectly elastic, nonisolable systems as isolated systems) is a pervasive part of model-building methodologies.

#### 4 Factual understanding

with other aims. Accurate prediction is achieved by tools poorly suited to explain; the aim of quick action is at odds with full causal representation; etc. At root, this is because the different aims of science are furthered by different means.

The idea that science has different, to some extent opposing aims, does a better job than truth-centered epistemology of accounting for the central role that idealizations play in scientific inquiry. For it regards idealizations as appropriate means to certain respectable ends – not as half-baked means to the ultimate goal of truth. To its further credit, the multi-purpose view of science also does justice to the plausible claim that idealizations do indeed distort, simplify, or overgeneralize in certain respects. The multi-purpose view recognizes idealizations as being indispensable to science, because they are necessary means to achieving some of science's central aims. But at the same time, it rightly recognizes them as being in some sense also defective, because their usefulness for certain purposes inevitably renders them less useful for other aims. A highly simplified model might be highly useful for predicting general trends, but it will, for that very reason, be ill-suited for predicting singular events. And a unified account of some historical period may contribute to the internal coherence of a narrative, whilst, by that token, fail to capture certain anomalous events within that period.

The account of factual understanding that I have developed dovetails with the view that science has diverse aims, and with the idea that the epistemic acceptability of idealizations can be accounted for in terms of a multi-purpose perspective on scientific inquiry. But my account is also more consistently pragmatist than some of the views proponents of the multi-purpose perspective have advanced, and for that reason preferable to those other views. For instance, Elgin, and to a lesser extent Potochnik, do continue to reserve a place for truth in their views on the epistemic value of idealization. Elgin (2004, p. 119), for instance, conceives of truth as a kind of 'threshold concept' (Potochnik. 2015, p. 73):

[I]n some cognitive endeavors we accept claims that we do not consider true. But we do not indiscriminately endorse falsehoods either. The question then is what makes a claim acceptable? Evidently, to accept a claim is not to take it to be true, but to take it that the claim's divergence from truth, if any, is negligible. The divergence need not be small, but whatever its magnitude, it can be safely neglected. We accept a claim, I suggest, when we consider it true enough. The success of our cognitive endeavors indicates that we are often right to do so. If so, a claim is acceptable when its divergence from truth is negligible. In that case it is true enough.

To my mind, the inclusion of a toned-down veridicality requirement in an otherwise pragmatist conception of science is problematic. The view according to which science has various (sometimes opposing) aims, fits well with actual scientific practice, licenses its many achievements, and makes sense of problems and dilemmas that scientific practitioners continuously face. But to insist that truth

operates alongside, or ‘underneath’ science’s various aims, threatens to undermine what is attractive about the pragmatist outlook. For truth, we have learnt from Peirce, is not an aim the achievement of which has identifiable consequences. To say that the products of science should be such that they sufficiently correspond to how the world really is, is to burden scientists with a requirement for the satisfaction of which they lack clearly defined standards of success and failure. Truth as an aim of science, even in its toned-down form, is a pragmatically spurious aim. Adding a veridicality requirement thus robs the pragmatist view of its internal coherence, and thereby makes it considerably less attractive.

The inclination on the part of some theorists to mitigate their pragmatism by coupling it with a soft veridicality requirement might stem from a lingering adherence to the standard, intuition-based method of conceptual analysis. And this ambivalence, in turn, might be due to a feeling that there is no immediate alternative to that method. The advantage of my view is that it can serve to strengthen the reluctant pragmatism that we find in much recent work in epistemology and philosophy of science, by rooting that pragmatism in a more general pragmatist philosophical methodology: the explication of concepts on the basis of the aims that the practices of applying those concepts are directed at. So rooted, those sympathetic to pragmatism have at their disposal additional arguments to ward off accusations of counter-intuitiveness.

### *3. The relation between objectual and explanatory understanding*

In section 1.1, I explained that grammatically different ways of ascribing understanding sometimes intuitively coincide with one another at an epistemological level. This is clearly the case for expressions of the form ‘to understand + noun (phrase)’ and expressions of the form ‘to understand why ...’. For instance, at least part of what constitutes having an understanding of natural phenomena consists in being able to explain those phenomena. Then again, there is also ample reason to think that explanatory understanding does not exhaust objectual understanding. For not all understanding of noun-denoted objects seems to come in the form of having, or being able to adduce, explanations pertaining to those objects: the understanding of symbols provides a case in point.<sup>140</sup> Recently, however, Kareem Khalifa (2017, p. 81), has argued that “aside from terminological convenience, anything noteworthy about objectual understanding can be replicated by explanatory understanding without loss”. Apart from its purely theoretical interest, the elucidation of the relation between objectual and explanatory understanding is bound to have broader ramifications for the philosophical study of understanding. Thus far, epistemologists and philosophers of science have tended to exploit unresolved ambiguities in establishing the primacy of either concept (i.e. objectual or explanatory

---

<sup>140</sup> For further arguments in favor of the non-equivalence of objectual and explanatory understanding, see, for instance, Kvanvig (2009) and Gijssbers (2013).



#### 4 Factual understanding

understanding) over the other. This has perpetuated a dialectic in which both ‘camps’ sometimes talk at cross-purposes. To illustrate this, I shall in what follows highlight two ways in which Khalifa’s influential position hinges on the (unintended?) exploitation of ambiguities, and suggest how my account of factual understanding might inspire a more transparent protocol for theory-building. My aim here is not to issue a definite verdict regarding the relation between objectual and explanatory understanding, but to point out how my account of factual understanding can help to remedy a counterproductive dialectic.

As said, Khalifa has defended a view according to which objectual understanding reduces to explanatory understanding. This view hinges, first of all, on a narrow conception of the extension of the term ‘objectual understanding’. For Khalifa, objectual understanding pertains only, or at least primarily, to the understanding of subject matters. So conceived, only the use of expressions like ‘Sandra understands organic chemistry’ and ‘Peter understands macroeconomics’ would qualify as ascriptions of objectual understanding. A second feature of Khalifa’s account is that it relegates ostensibly non-explanatory kinds of objectual understanding to the domain of explanatory understanding, by arguing that any epistemic value accruing to such non-explanatory understandings derives from their being instrumental to the acquisition of explanatory understanding. These two aspects of Khalifa’s view jointly conspire to make his reductionism more plausible, in the following ways. First of all, subject matters tend to be defined and demarcated within practices (i.e. scientific disciplines) which to a large extent exist in order to supply explanations. Khalifa’s restrictive conceptualization of objectual understanding thus helps to steer that concept in the direction of explanatory understanding. And secondly, Khalifa’s contention that the epistemic value of non-explanatory forms of objectual understanding hinges on their being instrumental to obtaining explanatory understanding, allows him to rebut a number of potential counterexamples which would appear to show that objectual understanding can bypass explanation, such as those discussed by Lipton (2009).<sup>141</sup>

My objective is not to discuss Khalifa’s arguments in detail and to take issue with those arguments directly. Rather, I want to point out that Khalifa’s arguments are made within, and rely for their cogency on, a dialectic situation that is marred by unresolved ambiguities. First of all, Khalifa’s choice for a restrictive conception of objectual understanding that is almost suspiciously well-tailored to his ulterior purposes, could only have been made in a dialectic context in which the extension

---

<sup>141</sup> Lipton’s (2009, p. 45) most well-known example is the following:

I never properly understood the why of retrograde motion until I saw it demonstrated visually in a planetarium. A physical model such as an orrery may do similar cognitive work. These visual devices may convey causal information without recourse to an explanation. And people who gain understanding in this way may not be left in a position to formulate an explanation that captures the same information. Yet their understanding is real.

of ‘objectual understanding’ is not yet fixed. The fact that epistemologists and philosophers of science have relied on diverging conceptions of objectual understanding, has created a kind of smokescreen behind which strategic choices regarding the meaning of that concept can be presented as harmless basic assumptions. Secondly, a similar lack of consensus regarding the extension of ‘explanatory understanding’ enables Khalifa (2017, p. 86) to subsume what he calls ‘proto-understanding’ under the heading of explanatory understanding. More specifically, Khalifa uses the lack of consensus to his advantage, by taking the liberty to conceptualize explanatory understanding in terms of *scientific* knowledge of an explanation. This relatively demanding and encompassing conceptualization lends credibility to the idea that someone who strictly falls short of having explanatory understanding may nonetheless be considered to be ‘on track’ to having that understanding. For if some epistemic standing is conceived of as being both wide-ranging and difficult to acquire, then room is created for arguing that ostensibly different epistemic standings are but stepping stones to the big prize.

All this is not to say that Khalifa’s arguments fail. But what I am trying to convey is that those arguments are predicated on rather consequential assumptions – assumptions that Khalifa is able to make more or less with impunity. The main reason why this intellectual wiggle room exists, is that in the present dialectic the extensions of key terms tend to be demarcated on the basis of intuitive judgments about the homogeneity of those extensions. Such judgments are often difficult to contend with, because they are intrinsically private and subject to variation among individuals. Some may deem the grammatically delineated conception of objectual understanding (‘to understand + noun (phrase)’) to capture an epistemologically uniform category, whilst others might be inclined to partition this grammatical conception further into distinct epistemological types, such as understanding subject matters, understanding phenomena, and understanding persons. Both options would seem *prima facie* legitimate. Problematically, however, the private and variable character of intuitions renders those intuitions amenable to being exploited for strategic purposes without this making the theorist liable to direct criticism. The inscrutability of intuitions confers plausible deniability upon any suspicion that those intuitions are selected and utilized for ulterior purposes.

My account of factual understanding provides a general protocol for fixing the extensions of core terms in a way that allows for choices so made to be critically engaged with. The general approach on which my account is built dictates that we first take a linguistic situation at face value, then draw intuition-based subdistinctions (if any), *and subsequently* make a reasoned case as to which practice, or practices, are to be associated with those intuitive categories. In this inquiry, I first isolated a certain grammatical category (objectual understanding), then made an intuitive distinction within that category (symbolic vs. factual understanding) and finally argued which purposes might be served by the ascription of these intuitive categories. The final step offers a check upon the second: the viability of an intuitive

#### 4 Factual understanding

distinction can be confirmed, but also undercut, by one's success or failure in characterizing the purposes associated with ascribing the supposedly distinct categories. Whether one so succeeds or fails depends on how strong the arguments are that are put forward in support of one's differentiation of the purposes. My three-step procedure is thus more transparent, and hence more open to critical discussion, than a two-step procedure in which there is no extra check upon our intuitive judgments.

I close off by indicating how we should go about assessing the relation between objectual and explanatory understanding, based on the protocol I have just sketched. To be sure, I will not commit myself to a view on which objectual and explanatory understanding are distinct, but only explain what a defense of any such view should look like. The first thing to note is that, according to my taxonomy, any overlap between objectual and explanatory understanding would be specifiable in terms of an overlap between factual and explanatory understanding. Symbolic understanding, after all, is clearly distinct from explanatory understanding. The more interesting question is thus whether factual and explanatory understanding are equivalent or not. The answer to this question hinges on which purposes one thinks are served by ascriptions of either. As I have argued, ascribing factual understanding serves the purpose of optimizing a community's attunement to its environment. One's degree of attunement depends on the extent of one's ability to achieve goals through engaging in efforts to achieve those goals. Intuitively, this goal is more general than the goal associated with ascribing explanatory understanding: the latter appears to be especially associated with the purpose of reliable prediction, whereas factual understanding may be thought to serve various other purposes besides prediction. However, given the aforementioned problems surrounding the reliance on intuitive judgments, it would not quite do to insist on a distinction between factual and explanatory understanding on such a basis. Moreover, what use would a community have in employing both concepts, if one of those concepts already covered for all the needs that were catered to by the other?

A more promising defense of the non-equivalence of factual and explanatory understanding takes the opposite direction, by first noting which needs a community might have in managing its epistemic endeavors. One plausible suggestion could be that a community may sometimes want to optimize the environmental attunement of a maximal proportion of individuals within a community, whilst at other times prioritize maximizing the attunement of certain privileged individuals. The former option 'ecumenizes' the epistemic situation within a community; the latter option promotes the division of epistemic labor. The two aims are not fully reconcilable: prioritizing the one over the other will likely come at the expense of the other to at least some extent. Given this irreconcilability, and assuming that both aims are of importance, there may be a need within a community to have two distinct concepts to see to it that both goals are met (albeit in different contexts). From this perspective, one could argue for a distinction between factual and explanatory understanding –

not on the basis of any intuitive difference between the two, but rather on the basis of an observation about the needs of a community. To be sure, I am not pledging allegiance to this view – I just wish to illustrate what kind of argument would have to figure in defending it.

#### *4. Knowledge, understanding, and scientific progress*

There is a broad consensus among epistemologists that knowledge requires truth. Importantly, also, many agree that “[t]ruth is a metaphysical, as opposed to epistemological, notion: truth is a matter of how things are, not how they can be shown to be. So when we say that only true things can be known, we’re not (yet) saying anything about how anyone can access the truth” (Ichikawa & Steup, 2018). Those who are committed to these standard views, may well be committed to the following, equally reasonable-sounding claims: (1) the accumulation of knowledge is the main, if not exclusive aim of science, and (2) scientific progress is measured in terms of the accumulation of knowledge. As Bird (2007, p. 64) rightly points out, the idea that “an episode in science is progressive when at the end of the episode there is more knowledge than at the beginning” has undeniable intuitive appeal. In particular, it nicely captures our sense that apparently fruitful episodes in scientific history which we later found to be completely misguided, did, in the end, not (or only marginally) contribute to scientific progress. In other words, the cumulative, knowledge-centered picture does justice to our intuitive, non-relativistic conception of progress. However well our solutions to scientific problems seem to work, they do not really constitute progress if they do not latch on to how the world really is.

However, the cumulative account of progress suffers from the same general problem that any truth-as-correspondence-centered epistemological theory is bound to suffer from: the fact that truth-as-correspondence cannot be a genuine purpose of our intellectual endeavors (see section 4.4, #4). For, if it is assumed that “truth is a matter of how things are, not how they can be shown to be”, then it is ultimately beyond our ken to ascertain under which conditions truth has or has not been arrived at. For this reason, science cannot genuinely aim for knowledge in the traditional sense (although scientists may of course believe they endeavor to gain knowledge), nor can scientific progress really be measured in terms of the accumulation of knowledge. That is to say, we could in principle continue to hold onto an ideal of science, and of scientific progress, as being centered around the accumulation of knowledge, and perhaps the mere entertainment of that ideal has certain beneficial implications for how scientific inquiry is conducted. But we cannot organize our scientific practices around those ideals. In influential work, Laudan (1977, p. 127) expresses this concern as follows:

[W]e apparently do not have any way of knowing for sure (even with some confidence) that science is true, or probable, or that it is getting closer to the truth. Such aims are *utopian*, in the literal sense that we can never know whether they are

#### 4 Factual understanding

being achieved. To set them up as goals for scientific inquiry may be noble and edifying to those who delight in the frustration of aspiring to that which they can never (know themselves to) attain; but they are not very helpful if our object is to explain how scientific theories are (or should be) evaluated.

If science is to be purposeful, it must be directed at purposes the achievement of which can be practically distinguished from their non-achievement. Consequently, since scientific progress is plausibly measured in terms of how well science lives up to its aims, the notion of scientific progress must likewise make reference to purposes that are genuine.

My explicatum of factual understanding has the advantage that it can provide for an account of the aim of science, and of the nature of scientific progress, around which our scientific practices can actually be organized. If the purpose of science is to increase our factual understanding (in my sense of that concept), then the extent to which science fulfills this purpose can be measured by finding out how strongly science contributes to the dissemination of useful information. Obviously, this requires that we answer a number of difficult questions about which types of information are most useful, and how scientists had best contribute to the dissemination of information. As attested to by the advent of Copernican astronomy (see, again section 4.4 #4), the purposes of a community tend to change as new measuring and observation devices, as well as new concepts and taxonomies, are invented. Those changing purposes bring about shifts in what information a community considers most useful. Furthermore, changing conceptions of how epistemic labor is best divided have ramifications for what constitutes the optimal protocol for disseminating information. Perhaps more so than at any previous stage in scientific history, contemporary science is geared towards specialization, which by and large favors the dissemination of esoteric information of high 'local' usefulness among limited numbers of individuals over the dissemination of more generally accessible information of inferior local usefulness within larger groups. Specialization gives rise to a proliferation of purposes, but tends to diminish the overall proportion of individuals within a community who can vouch for the achievement or non-achievement of those purposes.

My explicatum of factual understanding leaves room for different ideas about which types of purposes should be prioritized in aiming to optimize a community's environmental attunement, as well as for different ideas about whether environmental attunement is best promoted via a segregation- or an integration-oriented approach towards the dissemination of information. Verdicts about how well (if at all) science has progressed, and in what respects, thus hinge on how the notion of environmental attunement is cashed out in any particular day and age. This implies, as I pointed out before, a kind of diachronic relativism: there is no absolute measure of progress, but there are only context-dependent assessments of progress. To be sure, those context-dependent assessments may well show a certain continuity

over time. For instance, particular scientific developments might be so all-encompassing that they simultaneously cover all historical purposes as well as novel ones. Our tendency to look at scientific history as gradual absorption of previous purposes into more encompassing ones (think of how classical mechanics tends to be viewed as a limiting case of relativity theory) fosters linearity across consecutive assessments of progress. But the trajectory of gradual absorption + addition is not a given; our purposes can in principle evolve in any direction.

My account of scientific progress, which is based on my explicatum of factual understanding rather than on a traditional conception of knowledge, seems to have two significant problems. First, it makes short shrift of our intuitive, ‘birds-eye’ notion of progress – a notion that the knowledge-based approach does capture. Secondly, it appears to provide only for a general framework, in which many details are left to be filled in. Nonetheless, I deem my proposal to be preferable over the knowledge-based approach, given that the latter, in the end, does not even meet the minimal requirement of accounting for the aim of science. Without that requirement fulfilled, it cannot give a general story of what scientific progress consists in, let alone explain how assessments of progress are made in practice. I therefore deem the intuitive deficit of my account to be a comparatively less consequential shortcoming.<sup>142</sup>

##### 5. *The epistemic value problem*

Within epistemology, an important motivation for the turn to understanding stems from a desire to account for the distinctively epistemic value of our intellectual endeavors. The perennial question, which can be traced back to Plato’s *Meno*, is what value, if any, knowledge has over mere true belief. This so-called *primary value question* (Pritchard, 2007, p. 86) poses a problem for an important class of theories of knowledge – namely so-called reliabilist accounts of knowledge. According to reliabilism, knowledge amounts to reliably formed true belief. In other words, a true belief constitutes knowledge when it was formed on the basis of a truth-conducive belief-forming process. Zagzebski (2003) has argued that this view has trouble accounting for knowledge’s distinctive value. Pritchard, Turri & Carter (2018) reconstruct her argument as follows:

---

<sup>142</sup> Again, this aligns with Laudan’s (1977, p. 127) defense of his (similarly pragmatist) ‘problem-solving model’ of science:

The workability of the problem-solving model is its greatest virtue. In principle, we can determine whether a given theory does or does not solve a particular problem. In principle, we can determine whether our theories now solve more important problems than they did a generation or a century ago. If we have had to weaken our notions of rationality and progress in order to achieve this end, we are at least now in a position to be able to *decide* whether science is rational and progressive – a crucial necessity denied to us if we retain the classical connections between progress, rationality, and truth.

#### 4 Factual understanding

Zagzebski argues that the reliability of the process by which something is produced does not automatically add value to that thing, and thus that it cannot be assumed that the reliability of the process by which a true belief is produced will add value to that true belief. In defence of this claim, she offers the analogy of a cup of coffee. She claims that a good cup of coffee which is produced by a reliable coffee machine—i.e., one that regularly produces good cups of coffee—is of no more value than an equally good cup of coffee that is produced by an unreliable coffee machine.

Furthermore, as this line of objection goes, true belief is in the relevant respects like coffee: a true belief formed via a reliable belief-forming process is no more valuable than a true belief formed via an unreliable belief-forming process. In both cases, the value of the reliability of the process accrues in virtue of its tendency to produce a certain valuable effect (good coffee/true belief), but this means that where the effect has been produced—where one has a good cup of coffee or a true belief—then the value of the product is no greater for having been produced in a reliable way.

As the authors go on to explain, a possible way of resolving reliabilism's explanatory deficit could be to disavow standard process reliabilism and opt instead for a virtue-epistemological theory of knowledge. The benefit of such a theory is that it construes knowledge as incorporating "both the true belief and the source from which that true belief was acquired", and hence avoids being defeated by the coffee analogy.

Apart from the *primary value question*, however, epistemologists have been increasingly concerned with the *secondary value question* (Pritchard, 2007, p. 87), which asks what value, if any, knowledge has over any proper subset of its parts, including justified, but Gettierizable, true belief. Answering this question has proved difficult also for those theories which were in a position to answer the primary value question, such as virtue-epistemological accounts of knowledge. In fact, as explained already in section 1.2, it has been argued by Kvanvig (2003), that the prospects within Gettier-oriented epistemology for a satisfactory answer to the secondary value question are dim. Again, let me quote Pritchard, Turri & Carter (2018):

In essence, Kvanvig's argument rests on the assumption that it is essential to any virtue-theoretic account of knowledge—and any internalist account of knowledge as well, for that matter (i.e., an account that makes a subjective justification condition necessary for knowledge possession)—that it also includes an anti-Gettier condition. If this is right, then it follows that even if virtue epistemology has an answer to the primary value problem—and Kvanvig concedes that it does—it will not thereby have an answer to the secondary value problem since knowledge is not simply virtuous true belief. Moreover, Kvanvig argues that once we recognize what a gerrymandered notion a non-Gettierized account of knowledge is, it becomes apparent that there is nothing valuable about the anti-Gettier condition on knowledge that needs to be imposed. But if that is right, then it follows by even virtue epistemic lights that knowledge—i.e., non-Gettierized virtuous true believing—is no more valuable than one of its proper sub-sets—i.e., mere virtuous true believing.

If the study of understanding were to remain faithful to the knowledge-based paradigm associated with post-Gettier epistemology, then the problems posed by Kvanvig's argument would likely transfer to theories of understanding. This would be the case, for instance, if understanding were deemed to require anti-luck conditions that resemble or even match those associated with knowledge, as some authors have indeed argued.<sup>143</sup>

My account of factual understanding inherits few commitments from the post-Gettier tradition and is thus in a better position to offer a fresh solution to the epistemic value problem. Admittedly, it cannot directly answer the primary or secondary value questions – those questions were, after all, already stated in terms that presuppose (some aspect of) the tripartite analysis of knowledge. But it can answer the more general question after the value of factual understanding. What makes having factual understanding worthwhile? My account offers a deceptively simple, yet substantive answer: the value of factual understanding is implicated in the pragmatist outlook that undergirds my explication of that concept. That is, I have assumed that ascriptions of factual understanding serve a purpose to a community, and I have attempted to provide an account of factual understanding that makes sense of how ascriptions of factual understanding contribute to the achievement of that purpose. My answer to the epistemic value question is thus: assuming that the practice of ascribing factual understanding serves to promote behavior directed at the distribution of environmental information, and assuming also that ascriptions of factual understanding are by and large effective for realizing this aim, we can infer that factual understanding is valuable simply because its possession contributes to the achievement of an aim that communities have.

This answer might invite a further question, though: what value is there in having a practice that promotes the distribution of environmental information? The answer I already offered, namely that that value resides in the fact that the distribution of information helps to optimize a community's attunement to its environment, may not assuage everyone. For where is the value in *that*? Why would it be intrinsically worthwhile to strive to be more attuned to our environment? Unfortunately, there appears to be no necessary connection between being more attuned to our environment (in a pragmatic sense) and doing more good (in a moral sense). Oppenheimer possessed factual understanding by virtue of his disposition to distribute information useful to the development of the atomic bomb, but we would surely be reluctant to say that Oppenheimer's understanding was valuable because it directly contributed to the moral improvement of the human species. Perhaps, then, we should conclude that there is no deeper answer to the question what makes factual understanding worth having than that its possession serves a particular purpose – a purpose the realization of which might, sadly, jeopardize the achievement of more

---

<sup>143</sup> See DePaul & Grimm (2007) and Pritchard (2005) for defenses of understanding's (partial) incompatibility with luck.



#### 4 Factual understanding

lofty, moral purposes. This picture retains the traditional juxtaposition of ethics and epistemology, on which these disciplines are concerned with wholly disjoint forms of normativity.

I see one possible rebuttal to this line of thought – albeit a rather speculative one. Clearly, the idea that better factual understanding always runs in tandem with moral improvement is not very plausible. But there is some attraction to the idea that better factual understanding correlates with having a better epistemic position as regards matters of moral interest, and hence with greater *opportunity* for moral improvement. This is because better factual understanding implies not only increased success in realizing existing purposes, but often also (partial) success in realizing novel purposes. To at least some extent, then, having better factual understanding implies being able to do *more* things. The possibilities so unlocked are bound to be of varying moral commendability, ranging from the highly laudatory to the horribly repelling. As a result, having a broader range of alternatives for action will generally entail having to engage in more extensive moral deliberation. The more we are able to do, the bigger the set of moral considerations that we may have to go through in deciding what to do. Prior to the development of the atomic bomb, we did not need to think about the implications of possessing (let alone deploying) such a weapon of unfathomable destruction. The terrible fruits of the Manhattan project have alerted humanity to new types of consequences and new principles to take into account in moral decision-making. To be sure, I by no means want to say that humanity's overall moral standing has improved as a result of this. Indeed, that moral standing may well have deteriorated. But I would be inclined to say that humanity's *epistemic* position as regards matters of morality has improved (slightly) because of it.

Such an improved epistemic position is valuable because, in principle, it provides one with greater opportunity for moral improvement. After all, having access to a larger set of possible moral considerations that may impinge upon action, implies being in a position to act in accordance with more well-considered moral judgments. One could argue that better factual understanding is therefore an indirect means to moral improvement, and hence valuable not only in a purely practical sense, but also in the sense that it is instrumental to procuring something which is arguably of intrinsic value. I admit, however, that this instrumental relation is almost inevitably highly imperfect. For greater attunement to the environment also tends to implicate greater opportunity for moral degradation. As much as it may have improved our epistemic position as regards matters of morality, the development of the atomic bomb has opened up major new possibilities for morally reprehensible action. Indeed, these possibilities have materialized. The pursuit of factual understanding is thus risky business: it could well be that, in actuality, the increase of factual understanding will only ever beget more, and more consequential, morally disastrous behavior. The value of factual understanding, as I have argued here, is thus supposed to lie in its mere potential for improving our moral standing. So conceived, it is a *prima facie* value: any increase in factual understanding potentially contributes to

moral improvement, even if, in actuality, the increase of factual understanding only ever deepens our moral depravity.

#### **4.6 Conclusion**

In this chapter, I have formulated an explicatum of factual understanding and shown that this explicatum has considerable benefits over more traditional, intuition-based proposals regarding the nature of factual understanding. In essence, my explications of symbolic and factual understanding form a package deal: they are predicated on a joint methodological framework, and based on a shared set of assumptions. But the two elements of this package perform different functions within this epistemological manifesto: my explicatum of symbolic understanding has operated as a means to the explication of factual understanding; my explicatum of factual understanding has in turn served as a means for justifying (in Carnap's sense) the explicatum of symbolic understanding. Jointly considered, the two elements make for a thoroughly pragmatist contribution to the understanding literature. This places my proposal at some distance from what is common in contemporary epistemology, but it is a distance that I deem welcome given the fact that conventional approaches are not optimally equipped to deliver on the transformative vision that first motivated the turn to understanding.



## Concluding remarks

In conducting this investigation, I have employed a method which tells us to tailor our analysis of a concept to the purpose for which that concept is used. This method has been applied to two targets: symbolic and factual understanding. The resulting picture of these two concepts markedly differs from a picture that an intuition-based approach would have yielded. For as announced already in the introduction, my proposal is at odds with the commonsensical idea that symbolic and factual understanding consist, at least in part, in cognitive dispositions (e.g. beliefs). Instead, it says that both these concepts ultimately pertain to behavioral dispositions – with a couple of important qualifications. Underlying this is, one could say, a difference in perspective. The method of conceptual analysis allows us to retain our default, insider’s perspective on a practice we are ourselves immersed in. The method I have followed, by contrast, prompts us to take the slightly schizoid stance of an insider-looking-on-as-an-outsider. The latter approach, which admittedly comes less naturally than the former, impels one to view a certain practice as if one were a biological experimenter or an inquisitive extraterrestrial visitor. From this vantage point, the guiding question is not “what do I, in the capacity of user of concept  $x$ , take  $x$  to apply to?” but rather “what do I, in the capacity of observer of users of concept  $x$ , find  $x$  to be used for?”

Of course, just as intuitions may vary per theorist, assessments of purposes may differ depending on the observer. How I have characterized the purposes of ascribing symbolic and factual understanding may differ from someone else’s characterization of those purposes. But crucially, unlike diverging intuitive judgments, those different characterizations are up for intersubjective scrutiny: they can be objected to in terms of complaints of the following form: “Purpose  $y_1$  is not what concept  $x$  is used for. For given that the use of  $x$  consistently brings about a type of effect associated not with purpose  $y_1$  but with purpose  $y_2$ , and given that this consistency is due to behavioral variation on the part of users of  $x$ , we must conclude that  $x$ ’s purpose is not  $y_1$  but  $y_2$ .” As indicated in chapter 1, my account of the purposes of symbolic and factual understanding is based on ‘speculatively empirical’ hypotheses which are *prima facie* plausible, but which may in principle be undercut by empirical evidence. I grant that my proposals are therefore vulnerable to the type of complaint just mentioned.

One might think that this is a shortcoming of my approach. I think, however, that the immediate alternative is even less attractive. The theoretical stalemates that epistemologists have run into with respect to the analysis of knowledge, and which are liable to reoccur within the methodologically conservative study of understanding, are to an important extent due to the fact that theorists rely on subjective prioritizations of conflicting, but often equally legitimate intuitions. Eminently reasonable claims about the context-dependency of knowledge

## Concluding remarks

ascriptions run counter to likewise reasonable, context-independent concerns about the threat of skepticism. Similarly, many understandably harbor an internalist, achievement-related notion of understanding, which potentially conflicts with the demand for truth or truth approximation that understanding also seems to make. Some choice has to be made; some intuition must prevail over the other – lest aporia ensues. This means, however, that traditional epistemological debates are bound to gradually evolve into a more or less static co-existence of positions, where argumentative exchange often does more to further entrench already opposing camps than to stimulate the mutual reconsideration of each other's views. To be sure, the purpose-centered method is likely to lead to (persistent) disagreements as well. But given the comparatively stronger empirical element, there is at least more common ground to rely on, more of an independent adjudicator to appeal to.

Has this then, in the end, been a naturalistic manifesto? Yes and no. My explications are naturalistic in the sense that they subsume concept-use under the general heading of purposeful behavior, and thereby effectively categorize it as just another study object for ethologists. At the same time, many of the issues I have addressed cannot be reduced to scientific questions. For instance, my aim has not merely been to identify the purposes of the concepts of symbolic and factual understanding, but to explicate those concepts in a way that makes sense of how they are means to their purposes. This task involves genuine philosophical work, as there may be multiple, equally effective means to realizing a purpose. I have argued in this dissertation for particular conceptions of the means to realizing the purposes of maintaining communicative efficiency, and of optimizing environmental attunement. But it might be that, in light of those very same purposes, there are other ways of defining symbolic and factual understanding which make better sense of what we use those concepts for. Perhaps there are simpler ways, more exact ways, or more fruitful ways. Such matters are not settled solely on the basis of empirical data, but require the weighing of theoretical values.

In further deviation from (radical forms of) naturalism, there are some substantively normative questions associated with my purpose-centered approach which I have not addressed in this thesis – except for some brief comments in chapter 4. One category of questions has to do with the fact that a community may itself come to reflect upon what it uses its concepts for, and indeed upon what it *should* use its concepts for. Potential conflict of epistemic purposes with non-epistemic ones may occasion such reflection: which research should be funded, given the larger interests of the community? Should considerations of accessibility impinge upon how demanding our symbol-using practices are? Additional pressure to confront normative issues comes from the fact that certain innovations may change how we conceive of our goals. Environmental attunement in the general sense is a non-negotiable purpose, but there are increasingly many ways in which the specifics of this purpose can be thought of. Consider only the proliferation of artefacts which offer convenient alternatives to optimizing our attunement. Such developments bring

along normative questions. For example, should we delegate some, or even most, of our information-distributing activities to artefacts if that allows us to focus on other objectives? A similar question applies to the goal of maintaining communicative efficiency: to what extent, if at all, should we want to let artefacts do our symbol-usage for us if doing so reduces the overall effort and time needed to maintain communicative efficiency?

How we answer such questions may have drastic implications for what we take factual and symbolic understanding to consist in. Given the considerable distance that had to be bridged with the traditional approach to the study of understanding, my primary concern in this inquiry has been with answering the most general questions surrounding the nature of symbolic and factual understanding. It is therein that the main contribution of this thesis lies. In addition, however, I hope to have opened up a new framework for tackling some of the more specific issues just mentioned – a framework that is strongly pragmatist, moderately behaviorist and lightly naturalist.



## Nederlandse samenvatting

De epistemologie, of kenleer, heeft zich van oudsher beziggehouden met de vraag naar de aard van kennis. Een gangbare opvatting, traditioneel toegeschreven aan Plato, stelt dat kennis gelegen is in *overtuigingen* die *waar* en in epistemische zin *gerechtvaardigd* zijn. Nadat Edmund Gettier in 1963 zijn even beknopte als invloedrijke artikel ‘Is justified true belief knowledge?’ publiceerde, werd deze zogeheten ‘driedelige analyse’ onderwerp van verhitte discussie. Gettiers opvatting was dat die analyse tekortschoot: er zijn gevallen te bedenken – gevallen van *epistemisch toeval* – waarin aan de drie condities voldaan is, maar waarin intuïtief toch geen sprake is van kennis. De zoektocht naar een nieuwe definitie van kennis die weerbaar was tegen epistemisch toeval vond zijn weerslag in een almaar uitdijende en in toenemende mate technisch-filosofische literatuur, die zich richtte op het herzien of aanvullen van de driedelige analyse.

Aan het einde van de 20<sup>e</sup> eeuw kwamen enkele prominente epistemologen in het geweer tegen deze ontwikkeling. Auteurs als Catherine Elgin, Jonathan Kvanvig en Linda Zagzebski brachten uiteenlopende argumenten naar voren voor de stelling dat de door Gettierproblematiek bevangen epistemologie onvoldoende kon beantwoorden aan de vereisten die redelijkerwijs aan de kenleer gesteld mochten worden. Een gezonde epistemologie, aldus Elgin, Kvanvig en Zagzebski, geeft rekenschap van de verdiensten van de moderne (natuur)wetenschappen, verklaart de (meer)waarde van kennis, en onderkent, buiten de standaardnotie van overtuiging, het epistemische belang van cognitieve vermogens en gedragscompetenties. Om toekomstig kentheoretisch onderzoek in lijn te kunnen brengen met deze en andere doelstellingen, zouden epistemologen er goed aan doen hun discipline te herijken rond een nieuw concept: *begrip*.

Een aanzienlijk aantal kentheoretici bleek ontvankelijk voor deze oproep. Gewapend met het vertrouwde begrippenkader van de 20<sup>e</sup>-eeuwse epistemologie gingen zij het kersverse concept te lijf. Het beoogde nieuwe paradigma ging hierdoor echter al snel parallellen vertonen met de traditionele studie van kennis. In de recente literatuur komen die parallellen tot uiting in het type vragen dat opgeld doet, en in de bij de beantwoording van die vragen gehanteerde methode. Enkele regelmatig terugkerende vraagstukken zijn de volgende: veronderstelt begrip waarheid? Welke vorm van epistemische rechtvaardiging is van toepassing op begrip? En zelfs: is de notie van begrip gevoelig voor ondermijning door epistemisch toeval? Net als in de 20<sup>e</sup>-eeuwse discussies gaan epistemologen te rade bij hun intuïties om deze vragen van antwoorden te voorzien. De aloude *conceptuele analyse* blijft zodoende de vigerende methodiek.

In dit proefschrift merk ik allereerst op dat deze conservatieve aanpak vanuit methodologisch oogpunt problematisch is, en daarnaast onvoldoende tegemoetkomt aan de bezwaren en ambities van de wegbereiders van de kentheoretische studie van



begrip. Het methodologische probleem is gelegen in het feit dat, anders dan bij het min of meer binaire kennisconcept, analyses van het concept begrip zich vanwege het inherent graduele karakter van begrip maar matig lenen voor toetsing aan intuïties. De meer fundamentele zorg is dat de behoudende benadering leidt tot een studie van begrip die te zeer schatplichtig is aan het traditionele paradigma om de door Elgin, Zagzebski, Kvanvig, e.a. benoemde problemen van het post-Gettiertijdperk het hoofd te kunnen bieden.

Vertrekkend vanuit deze twee constateringën introduceer ik in **hoofdstuk 1** een alternatieve zienswijze. Ik richt me hierbij op twee vormen van begrip: enerzijds op die vorm die betrekking heeft op symbolen (woorden, zinnen, diagrammen, etc.) en anderzijds op het begrip dat we kunnen hebben van voorwerpen, gebeurtenissen en (natuur)fenomenen in de wereld. Ik betoog dat een studie van deze twee concepten – *symboolbegrip* en *feitbegrip* – het meest gebaat is bij de methode van *Carnapiaanse explicatie*. Deze methode is geënt op een viertal criteria, waarvan *vruchtbaarheid* het belangrijkste is. Binnen het kader van dit onderzoek geef ik als volgt invulling aan dit criterium: de voornoemde concepten dienen zo geëxpliciteerd te worden dat aan de hand van de resulterende *explicata* antwoorden gegeven kunnen worden op onder meer die probleemstellingen die in eerste aanleg aanleiding gaven tot de studie van begrip. Het voornaamste middel om dit doel te bereiken bestaat eruit om bij de explicatie van symbool- en feitbegrip uit te gaan van de *functies* die toeschrijvingen van deze vormen van begrip hebben binnen een bepaalde gemeenschap. Primair heeft dit proefschrift dus tot doel om de concepten symbool- en feitbegrip te expliciteren op grond van hun functies, en wel zodanig dat daarmee nieuw licht geworpen kan worden op enkele prangende kwesties binnen de hedendaagse studie van begrip.

**Hoofdstuk 2** verricht het voorwerk voor de explicatie van het concept symboolbegrip. Dit voorwerk behelst het definiëren van de term ‘symbool’. Ik wend me tot de binnen de analytische filosofie maar spaarzaam ontgonnen semiotiek van Charles Sanders Peirce om me van deze taak te kwijten. Meer in het bijzonder maak ik gebruik van een reconstructie van Peirce-deskundige Thomas Short, die heeft laten zien hoe Peirce’ veelomvattende en complexe symbooltheorie nauw samenhangt met diens ideeën over doelmatigheid en doelmatig gedrag. De keuze voor deze invalshoek stelt me ten eerste in staat om de term ‘symbool’ van een brede, maar toch precieze definitie te voorzien. Die brede en precieze definitie is bedoeld om het concept symboolbegrip later op vruchtbare wijze te kunnen expliciteren. Ten tweede biedt Peirce’ theorie van doelmatigheid, die zijn semiotiek schraagt, de mogelijkheid om teleologische termen als ‘functie’, ‘doel’ en ‘handelen’ te ontdoen van hun problematische mentalistische lading.

In **hoofdstuk 3** maak ik mijn conclusies uit hoofdstuk 2 te gelde door ze in te zetten in de explicatie van symboolbegrip. Alvorens het eigenlijke explicatum te formuleren, beargumenteer ik dat het concept symboolbegrip, gezien de functie die toeschrijvingen van dat concept hebben, op behavioristische wijze uitgelegd moet

worden. Toeschrijvingen van symboolbegrip zien namelijk op het handhaven en stimuleren van bepaalde gedragsdisposities, te weten die gedragsdisposities die bijdragen aan een uniforme en stabiele communicatiepraktijk binnen een gemeenschap. In het licht van deze bevinding moet symboolbegrip opgevat worden in termen van *symboolgebruik* – meer specifiek *competent* symboolgebruik. Ik stel vast dat een competent symboolgebruiker iemand is die een op deugdelijke leerprocessen gestoelde dispositie heeft om een symbool te gebruiken conform de interpretatieregel die aan dat symbool ten grondslag ligt. In de afsluitende twee secties onderscheid ik ten slotte twee dimensies van gradualiteit van symboolbegrip: breedte en diepte.

Het in hoofdstuk 3 geformuleerde explicatum van symboolbegrip kan beschouwd worden als een op zichzelf staand resultaat van dit onderzoek. Niettemin vervult het binnen het bestek van dit proefschrift in de eerste plaats een instrumentele rol: die van noodzakelijk onderdeel binnen een vruchtbaar explicatum van feitbegrip. In **hoofdstuk 4** wordt dat explicatum uitgewerkt en onderbouwd. De structuur van het hoofdstuk weerspiegelt die van hoofdstuk 3. Allereerst constateer ik dat feitbegrip net als symboolbegrip betrekking heeft op gedrag, omdat toeschrijvingen van feitbegrip dienen ter bevordering van bepaalde gedragsdisposities. Het gaat hierbij om gedrag dat gericht is op het succesvol informeren van anderen omtrent hun omgeving (in brede zin). Ik beargumenteer dat feitbegrip zo beschouwd symboolbegrip vereist, en specificer vervolgens de notie van succesvol informeren in termen van het in staat stellen, middels symboolgebruik, van een ontvanger tot het verwezenlijken van een aan die ontvanger toegeschreven doel. In de laatste drie secties breng ik enkele resterende onduidelijkheden tot klaarheid, weerleg ik mogelijke bezwaren, en benoem ik de voordelen van het gepresenteerde explicatum van feitbegrip. Ik concludeer dat de in dit proefschrift geformuleerde explicata van symbool- en feitbegrip tezamen de voorkeur verdienen boven op intuïtieve oordelen gegrondveste interpretaties van die concepten, omdat

- (1) De explicata een meer coherente lezing bieden van de relatie tussen symbool- en feitbegrip.
- (2) Er met behulp van de explicata beter rekenschap gegeven kan worden van de epistemische voor- en nadelen van (wetenschappelijke) idealisaties.
- (3) De explicata aanleiding geven tot een minder arbitraire opvatting van de relatie tussen *voorwerpeijk* en *verklarend* begrip.
- (4) Op grond van de explicata een aannemelijker beeld ontstaat van de aard van wetenschappelijke vooruitgang.
- (5) De explicata beter toegerust zijn op de beantwoording van de vraag naar de (meer)waarde van begrip.

Het lijkt geen twijfel dat de in dit proefschrift verdedigde duale theorie van symbool- en feitbegrip enkele controversiële aspecten kent. Ik denk hierbij met name aan de

#### Nederlandse samenvatting

uitgesproken behavioristische en pragmatistische stellingnames. Het is immers duidelijk dat die lijnrecht indruisen tegen intuïties die ons doen neigen naar een conceptualisatie van begrip als iets mentaals, en als iets wat waarheid of tenminste waarheidsbenadering vereist. Gegeven de weinig hoopgevende vooruitzichten van de eerder geschetste conservatieve studie van begrip, en in het licht van de oorspronkelijke uitgangspunten achter die studie, beschouw ik het intuïtief deficit echter zonder meer als een aanvaardbaar en zelfs toe te juichen verlies. De dankzij het intuïtieve tekort gegenereerde voordelen scheppen mijns inziens namelijk een vruchtbare voedingsbodem voor een studie van begrip die een volwaardig alternatief vormt voor de traditionele studie van kennis.

## References

- Adams, F. & Clarke, M. (2005). Resurrecting the tracking theories. *Australasian Journal of Philosophy*, 83(2), 207-221.
- Anscombe, G.E.M. (1957). *Intention*. Ithaca (NY): Cornell University Press.
- Augustine. (1975). *De dialectica*. J. Pinborg (Ed.) and B.D. Jackson (Trans.). Dordrecht: Reidel.
- Bater, L.R. & Jordan, S.S. (2019). Selective attention. In V. Zeigler-Hill & T. Shackelford (Eds.), *Encyclopedia of personality and individual differences*. Cham: Springer. [https://doi.org/10.1007/978-3-319-28099-8\\_1904-1](https://doi.org/10.1007/978-3-319-28099-8_1904-1)
- Baumberger, C. (2014). Types of understanding: Their nature and their relation to knowledge. *Conceptus*, 40(98), 67-88.
- Baumberger, C. (2019). Explicating objectual understanding: Taking degrees seriously. *Journal for General Philosophy of Science*, 50, 367-388.
- Baumberger, C., Beisbart, C., Brun, G. (2017). What is understanding? In S.R. Grimm, C. Baumberger & S. Ammon (Eds.), *Explaining understanding: New perspectives from epistemology and philosophy of science* (pp. 1-34). New York: Routledge.
- Baumberger, C. & Brun, G. (2017). Dimensions of objectual understanding. In S.R. Grimm, C. Baumberger & S. Ammon (Eds.), *Explaining understanding: New perspectives from epistemology and philosophy of science* (pp. 165-189). New York: Routledge.
- Bird, A. (2007). What is scientific progress? *Noûs*, 41(1), 64-89.
- Broadbent, D. (1958). *Perception and communication*. London: Pergamon Press.
- Brun, G. (2016). Explication as a method of conceptual re-engineering. *Erkenntnis*, 81, 1211-1241.
- Bruner, J.S. (1956). *A study of thinking*. New York: Wiley.
- Campbell, C.A. (1967). Towards a definition of belief. *Philosophical Quarterly*, 17(68), 204-220.

## References

- Carnap, R. (1950). *Logical foundations of probability*. Chicago: University of Chicago Press.
- Carter, J.A., Pritchard, D., & Shepherd, J. (2019). Knowledge-how, understanding-why and epistemic luck: An experimental study. *Review of Philosophy and Psychology*, 10, 701-734.
- Chandler, D. (2017). *Semiotics: The basics*. London: Routledge.
- Craig, E.L. (1990). *Knowledge and the state of nature*. Oxford: Clarendon Press.
- Cruse, A. (2000). *Meaning in language*. Oxford: Oxford University Press.
- Dennett, D.C. (1987). *The intentional stance*. Cambridge (MA): MIT Press.
- DePaul, M. & Grimm, S.R. (2007). Review essay: Kvanvig's The value of knowledge and the pursuit of understanding. *Philosophy and Phenomenological Research*, 74, 498-514.
- De Regt, H.W. (2009). The epistemic value of understanding. *Philosophy of Science*, 76, 585-597.
- De Regt, H.W. (2017). *Understanding scientific understanding*. Oxford: Oxford University Press.
- De Regt, H.W. & Gijsbers, V. (2017). How false theories can yield genuine understanding. In S.R. Grimm, C. Baumberger & S. Ammon (Eds.), *Explaining understanding: New perspectives from epistemology and philosophy of science* (pp. 50-75). New York: Routledge.
- DeRose, K. (1995). Solving the skeptical problem. *Philosophical Review*, 104(1), 1-52.
- Deutsch, J.A. & Deutsch, O. (1963). Attention: some theoretical considerations. *Psychological Review*, 70, 80-90.
- Dewulf, F. (2018). *A genealogy of scientific explanation: The emergence of the deductive-nomological model at the intersection of German historical and scientific philosophy* [Doctoral dissertation]. University of Gent. Retrieved from: <http://hdl.handle.net/1854/LU-8562228>.

- Dutilh Novaes, C. (2020). Carnapian explication and ameliorative analysis: A systematic comparison. *Synthese*, *197*, 1011-1034.
- Dutilh Novaes, C. & Reck, E. (2017). Carnapian explication, formalisms as cognitive tools, and the paradox of adequate formalization. *Synthese*, *194*, 195-215.
- Elgin, C.Z. (1996). *Considered judgment*. Princeton (NJ): Princeton University Press.
- Elgin, C.Z. (2004). True enough. *Philosophical Issues*, *14*, 113-131.
- Elgin, C.Z. (2009). Appendix C: Is understanding factive? In A. Haddock, A. Millar & D. Pritchard (Eds.), *Epistemic value* (pp. 322-330). Oxford: Oxford University Press.
- Elgin, C.Z. (2017). *True enough*. Cambridge (MA): MIT Press.
- Elliott, K.C. & McKaughan, D.J. (2014). Nonepistemic values and the multiple goals of science. *Philosophy of Science*, *81*, 1-21.
- Fetzer, J. (2017). Carl Hempel. In E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from: <https://plato.stanford.edu/entries/hempel>.
- Frankfurt, H.G. (1978). The problem of action. *American Philosophical Quarterly*, *15*(2), 157-162.
- Friedman, M. (1974). Explanation and scientific understanding. *Journal of Philosophy*, *71*(1), 5-19.
- Gettier, E.L. (1963). Is justified true belief knowledge? *Analysis*, *23*(6), 121-123.
- Gijsbers, V. (2013). Understanding, explanation, and unification. *Studies in History and Philosophy of Science*, *44*(3), 516-522.
- Glüer, K. & Wikforss, Å. (2015). Meaning normativism: Against the simple argument. *Organon F*, *22*(Supplementary issue), 63-73.
- Goldman, A.I. (1967). A causal theory of knowing. *Journal of Philosophy*, *64*(12), 357-372.
- Goodman, N. (1968). *Languages of art*. Indianapolis: Bobbs-Merrill.
- Goodman, N. (1978). *Ways of worldmaking*. Indianapolis: Hackett.

## References

- Gordon, E.C. (2012). Is there propositional understanding? *Logos & Episteme*, *III*(2), 181-192.
- Greco, J. (2014). Episteme: Knowledge and understanding. In K. Timpe & C.A. Boyd (Eds.), *Virtues and their vices* (pp. 285-302). Oxford: Oxford University Press.
- Grimm, S.R. (2010). Understanding. In S. Bernecker & D. Pritchard (Eds.), *The Routledge companion to epistemology* (pp. 84-94). New York: Routledge.
- Grimm, S.R. (2014). Understanding as knowledge of causes. In A. Fairweather (Ed.), *Virtue epistemology naturalized: Bridges between virtue epistemology and philosophy of science* (pp. 329-345). Cham: Springer.
- Grimm, S.R. (2017). Understanding and transparency. In S.R. Grimm, C. Baumberger & S. Ammon (Eds.), *Explaining understanding: New perspectives from epistemology and philosophy of science* (pp. 212-229). New York: Routledge.
- Grimm, S.R, Baumberger, C. & Ammon, S. (Eds.). (2017). *Explaining understanding: New perspectives from epistemology and philosophy of science*. New York: Routledge.
- Hempel, C.G. (1965). *Aspects of scientific explanation*. New York: Free Press.
- Hempel, C.G. (1966). *Philosophy of natural science*. Englewood-Cliffs (NJ): Prentice Hall.
- Hempel, C.G. & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, *15*(2), 135-175.
- Hendricks, V.F. (2006). *Mainstream and formal epistemology*. Cambridge: Cambridge University Press.
- Hetherington, S. (2001). *Good knowledge, bad knowledge*. Oxford: Oxford University Press.
- Horwich, P. (1998). *Meaning*. Oxford: Oxford University Press.
- Ichikawa, J.J. & Steup, M. (2018). The analysis of knowledge. In E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from: <https://plato.stanford.edu/entries/knowledge-analysis>.

- Kelp, C. (2015). Understanding phenomena. *Synthese*, 192, 3799-3816.
- Kelp, C. (2017). Towards a knowledge-based account of understanding. In S.R. Grimm, C. Baumberger & S. Ammon (Eds.), *Explaining understanding: New perspectives from epistemology and philosophy of science* (pp. 251-271). New York: Routledge
- Khalifa, K. (2017). *Understanding, explanation, and scientific knowledge*. Cambridge: Cambridge University Press.
- Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In P. Kitcher & W.C. Salmon (Eds.), *Scientific explanation* (pp. 410-505). Minneapolis: University of Minnesota Press.
- Kruschke, J.K. (2005). Category learning. In K. Lamberts & R.L. Goldstone (Eds.), *Handbook of cognition* (pp. 183-201). London: SAGE Publications.
- Kuhn, T.S. (1977a). Objectivity, value judgment and theory choice. In T.S. Kuhn (Ed.), *The essential tension* (pp. 329-339). Chicago: University of Chicago Press.
- Kuhn, T.S. (1977b). Second thoughts on paradigms. In T.S. Kuhn (Ed.), *The essential tension* (pp. 293-319). Chicago: University of Chicago Press. Reprinted from *The structure of scientific theories* (pp. 459-482), F. Suppe (Ed.), 1974, Urbana: University of Illinois Press.
- Kvanvig, J.L. (2003). *The value of knowledge and the pursuit of understanding*. New York: Cambridge University Press.
- Kvanvig, J.L. (2009). The value of understanding. In A. Haddock, A. Millar & D. Pritchard (Eds.), *Epistemic value* (pp. 95-111). Oxford: Oxford University Press.
- Kvanvig, J.L. (2010). Sosa's virtue epistemology. *Crítica*, 42(125), 47-62.
- Kvanvig, J.L. (2017). Understanding. In W.J. Abraham & F.D. Aquino (Eds.), *The Oxford handbook of the epistemology of theology* (pp. 175-190). Oxford: Oxford University Press.
- Laudan, L. (1977). *Progress and its problems: Towards a theory of scientific growth*. Berkeley: University of California Press.



## References

- Lear, J. (1988). *Aristotle: The desire to understand*. Cambridge: Cambridge University Press.
- Lehrer, K. & Paxson, T. (1969). Knowledge: Undefeated justified true belief. *Journal of Philosophy*, 66(8), 225-237.
- Leitgeb, H. & Carus, A. (2021). Rudolf Carnap. In E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from: <https://plato.stanford.edu/archives/sum2021/entries/carnap/>.
- Lewis, C.I. (1929). *Mind and the world-order: Outline of a theory of knowledge*. New York: Charles Scribner's Sons.
- Lewis, D. (1996). Elusive knowledge. *Australasian Journal of Philosophy*, 74(4), 549–567.
- Lipton, P. (2009). Understanding without explanation. In H.W. de Regt, S. Leonelli & K. Eigner (Eds.), *Scientific understanding: Philosophical perspectives* (pp. 43-63). Pittsburgh: University of Pittsburgh Press.
- Liu, D. (2008). Intransitive or object-deleting? Classifying English verbs used without an object. *Journal of English Linguistics*, 36(4), 289-313.
- Markus, R.A. (1957). St. Augustine on signs. *Phronesis*, 2(1), 60-83.
- Mayr, E. (1974). Teleological and teleonomic: A new analysis. In R.S. Cohen & M.W. Wartofsky (Eds.), *Methodological and historical essays in the natural and social sciences* (pp. 91-117). Dordrecht: Springer.
- McGinn, M. (2013). *The Routledge guidebook to Wittgenstein's Philosophical Investigations*. London: Routledge.
- McLeod, S.A. (2018). *Selective attention*. Retrieved December 1, 2021, from: <https://www.simplypsychology.org/attention-models.html>.
- Meier-Oeser, S. (2011). Medieval semiotics. In E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from: <https://plato.stanford.edu/archives/sum2011/entries/semiotics-medieval>.
- Millikan, R.G. (2004). *Varieties of meaning*. Cambridge (MA): MIT Press.

- Misak, C.J. (1991). *Truth and the end of inquiry: A Peircean account of truth*. Oxford: Clarendon Press.
- Moravcsik, J.M. (1979). Understanding and knowledge in Plato's philosophy. *Neue Hefte für Philosophie*, 15, 53-69.
- Moss, S. (2018). *Probabilistic knowledge*. Oxford: Oxford University Press.
- Murphy, G., Groeger, J. A., & Greene, C. M. (2016). Twenty years of load theory – Where are we now, and where should we go next? *Psychonomics Bulletin and Review*, 23(5), 1316-1340.
- Norricks, N.R. (1981). *Semiotic principles in semantic theory*. Amsterdam: John Benjamins Publishing Company.
- Nozick, R. (1981). *Philosophical explanations*. Cambridge (MA): Harvard University Press.
- Nunberg, G., Sag, I.A. & Wasow, T. (1994). Idioms. *Language*, 70(3), 491-538.
- Peijnenburg, J. (2000). Akrasia, dispositions and degrees. *Erkenntnis*, 53(3), 285-308.
- Peirce, C.S. (1931–58). *The collected papers of Charles Sanders Peirce*. C. Hartshorne & P. Weiss (Eds. Vols. 1-6), A. Burks (Ed. Vol 7-8). Cambridge (MA): Harvard University Press.
- Peirce, C.S. (1998). *The essential Peirce: Selected philosophical writings* (Vol. 2, 1893-1913). The Peirce Edition Project (Ed.). Bloomington (IN): Indiana University Press
- Plantinga, A. (1993). *Warrant: The current debate*. New York: Oxford University Press.
- Plato. (2015). *Plato: Theaetetus and Sophist*. C. Rowe (Ed. and transl.) Cambridge: Cambridge University Press.
- Potochnik, A. (2015). The diverse aims of science. *Studies in History and Philosophy of Science*, 53, 71-80.
- Potochnik, A. (2017). *Idealization and the aims of science*. Chicago: University of Chicago Press.

## References

- Potochnik, A. (2020). Idealization and many aims. *Philosophy of Science*, 87, 933-943.
- Pritchard, D. (2005). *Epistemic luck*. Oxford: Oxford University Press.
- Pritchard, D. (2007). Recent work on epistemic value. *American Philosophical Quarterly*, 44(2), 85-110.
- Pritchard, D. (2010). Understanding. In D. Pritchard, A. Millar & A. Haddock (Eds.), *The nature and value of knowledge: Three investigations* (pp. 66-88). Oxford: Oxford University Press.
- Pritchard, D. (2014). Knowledge and understanding. In A. Fairweather (Ed.), *Virtue epistemology naturalized: Bridges between virtue epistemology and philosophy of science* (pp. 315-328). Dordrecht: Springer.
- Pritchard, D., Turri, J. & Carter, J.A. (2018). The value of knowledge. In E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from: <https://plato.stanford.edu/archives/spr2018/entries/knowledge-value>.
- Putnam, H. (1975). *Mathematics, matter and method*. Cambridge: Cambridge University Press.
- Reichenbach, H. (1956). *The direction of time*. Berkeley: University of California Press.
- Riggs, W. (2003). Understanding virtue and the virtue of understanding. In M. DePaul and L. Zagzebski (Eds.), *Intellectual virtue: Perspectives from ethics and epistemology*. Oxford: Oxford University Press.
- Roberts, M.R. (2014, May 8). *What everybody gets wrong about Alanis Morissette's "Ironic"*. Salon. Retrieved from: [https://www.salon.com/2014/05/08/what\\_everybody\\_gets\\_wrong\\_about\\_alanis\\_morissettes\\_ironic\\_partner/](https://www.salon.com/2014/05/08/what_everybody_gets_wrong_about_alanis_morissettes_ironic_partner/)
- Rose, D. & Schaffer, J. (2013). Knowledge entails dispositional belief. *Philosophical Studies*, 166, 19-50.
- Ross, B.H. & Makin, V.S. (1999). Prototype versus exemplar models in cognition. In R.J. Sternberg (Ed.), *The nature of cognition* (pp. 205-241). Cambridge (MA): MIT Press.

Russell, B. (1948). *Human knowledge: Its scope and limits*. London: George Allen & Unwin.

Salmon, W.C. (1998). *Causality and explanation*. New York: Oxford University Press.

Short, T.L. (1981a). Peirce's concept of final causation. *Transactions of the Charles S. Peirce Society*, 17(4), 369-382.

Short, T.L. (1981b). Semeiosis and intentionality. *Transactions of the Charles S. Peirce Society*, 17(3), 197-223.

Short, T.L. (1983). Teleology in nature. *American Philosophical Quarterly*, 20(4), 311-320.

Short, T.L. (2007). *Peirce's theory of signs*. Cambridge: Cambridge University Press.

Sliwa, P. (2015). Understanding and knowing. *Proceedings of the Aristotelian Society*, 115(1), 57-74.

Smith, E.E. & Medin, D.L. (1981). *Categories and concepts*. Cambridge (MA): Harvard University Press.

Sober, E. (1984). *The nature of selection*. Cambridge (MA): MIT Press.

Sosa, E. (2007). *A virtue epistemology: Apt belief and reflective knowledge* (Vol. I). New York: Oxford University Press.

Stout, R. (2005). *Action*. Stocksfield: Acumen.

Strevens, M. (2008). *Depth: An account of scientific explanation*. Cambridge (MA): Harvard University Press.

Strevens, M. (2013). No understanding without explanation. *Studies in History and Philosophy of Science*, 44, 510-515.

Strevens, M. (2017). How idealizations provide understanding. In S.R. Grimm, C. Baumberger & S. Ammon (Eds.), *Explaining understanding: New perspectives from epistemology and philosophy of science* (pp. 37-49). New York: Routledge

## References

Tiercelin, C. (2006). The importance of the medievals in the constitution of Peirce's semeiotic and thought-sign theory. In R. Fabbrichesi & S. Marietti (Eds.), *Semiotics and philosophy in Charles Sanders Peirce* (pp. 158-184). Newcastle upon Tyne: Cambridge Scholars.

Tomasello, M. (2008). *Origins of human communication*. Cambridge (MA): MIT Press.

Turri, J., Alfano, M., & Greco, J. (2019). Virtue epistemology. In E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from: <https://plato.stanford.edu/archives/fall2019/entries/epistemology-virtue>.

Vicente, A. & Falkum, I.L. (2017). Polysemy. In M. Aronoff (Ed.), *Oxford research encyclopaedia of linguistics*. Retrieved from: <https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-325>.

Weslake, B. (2010). Explanatory depth. *Philosophy of Science*, 77(2), 273-294.

Wetzel, L. (2018). Types and tokens. In E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from: <https://plato.stanford.edu/archives/fall2018/entries/types-tokens>.

Whittlesea, B.W.A. (1987). Preservation of specific experiences in the representation of general knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(1), 3-17.

Wikforss, Å. (2001). Semantic normativity. *Philosophical Studies*, 102, 203-226.

Wilkenfeld, D.A. (2013). Understanding as representation manipulability. *Synthese*, 190(6), 997-1016.

Wilkenfeld, D.A. (2017). MUDdy understanding. *Synthese*, 194, 1273-1293.

Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.

Wittgenstein, L. (1953). *Philosophical investigations*. G.E.M. Anscombe & R. Rhees (Eds.), G.E.M. Anscombe (Trans.). Oxford: Blackwell.

Zagzebski, L.T. (1994). The inescapability of Gettier problems. *Philosophical Quarterly*, 44(174), 65-73.

## References

- Zagzebski, L.T. (1996). *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge*. New York: Cambridge University Press.
- Zagzebski, L.T. (2001). Recovering understanding. In M. Steup (Ed.), *Knowledge, truth, and duty: Essays on epistemic justification, responsibility, and virtue* (pp. 235–252). New York: Oxford University Press.
- Zagzebski, L.T. (2003). The search for the source of epistemic good. *Metaphilosophy*, 34(1-2), 12-28.
- Zagzebski, L. (2009). *On epistemology*. Belmont (CA): Wadsworth Cengage Learning.
- Zagzebski, L.T. (2019). Toward a theory of understanding. In S.R. Grimm (Ed.) *Varieties of understanding: New perspectives from philosophy, psychology, and theology* (pp. 123-136). New York: Oxford University Press.