

CHAPTER I

The Philosophy of Understanding

We seek to understand a great many things – the heights of the heavens; the depths of the oceans; motions and emotions; flora and fauna; friends, family, and foes; minds and machines; markets and morals; and much more. But understanding is only faintly understood. When do we limn the deep structures of the natural and social world? When do our attempts at grasping these structures misfire?

Understanding can be understood in different ways. An *empirical* inquiry might use experiments, interviews, surveys, and the like to uncover the mechanisms that make understanding possible, or it might track the historical arc of our various conceptions of understanding. By contrast, I will be pursuing a *theoretical* inquiry, in which I use various formal and conceptual apparatuses to construct a model of understanding.

As a *philosophical* theorist of understanding, my instruments are the tools of my specific trade. Thus, while the interesting theoretical concepts found in the cognitive and social sciences undoubtedly illuminate various aspects of understanding, my preferred instruments come from epistemology and the philosophy of science. Recent philosophical work on understanding straddles these venerable fields – crisscrossing with larger questions about justification, knowledge, and other cognitive achievements; the point and purpose of scientific explanation; the role of models and idealizations in scientific theories; and the pragmatic aspects of scientific inquiry. These and other philosophical forays provide a repository of ways to plumb the hidden depths of understanding.

Ideally, lively interdisciplinary crosstalk at both the empirical and theoretical levels should be achieved. Nevertheless, given the nascent state of the field, I hope that I will be forgiven for attempting to slay only a small handful of understanding's many dragons. And, lest I seem too indifferent to other areas of study, I will frequently argue that attention to scientific practice is an effective cure for the overzealousness that infects philosophy's more speculative organs.

1.1. Kinds of Understanding

Even among philosophical accounts of understanding, there are many kinds that can be studied and scrutinized. I will chiefly be interested in *explanatory* understanding (i.e., the understanding characteristic of good explanations). Paradigmatically, such understanding can be expressed as a kind of understanding-why; for instance, when we say, “Susan understands why the sky is blue.”

For ease of locution, I will treat “explanatory understanding” and “understanding-why” as synonyms. Having said this, explanatory understanding can be expressed without a “why.” For instance, I take my account to cover nearby examples, such as “Susan understands what causes the sky to be blue” and “Susan understands how it is that the sky is blue.”

Additionally, I will be restricting myself to understanding of *empirical* phenomena. I make no claims about understanding in ethics, aesthetics, mathematics, or logic, for example. This assumes, of course, that these are not empirical domains. My hankerings for a panoptic empiricism notwithstanding, I will stake no claim on these tangled conceptual thickets.

While other kinds of understanding, as listed in Table 1.1, are also worthy of study, they will not be my focus:

Table 1.1: *Different Kinds of Non-Explanatory Understanding*

Kind of understanding	Typical Complement	Examples
Propositional	<i>that</i> + declarative sentence	I understand that you might not enjoy reading this book.
Broad Linguistic	The name of a language	Schatzi understands German.
Narrow Linguistic	<i>what</i> + a linguistic expression + <i>means</i>	Schatzi understands what “Ich bin Berliner” means.
Procedural	<i>how</i> + infinitive	Miles understands how to play trumpet.
Non-explanatory Interrogative	Embedded question that does not seek an explanation as its answer (most who, where, what, and when questions)	I understand who my friends are. I understand where my friends will be going. I understand what my friends are doing. I understand when my friends need a good laugh.

1.2. Degrees of Understanding

3

The list in Table 1.1 is meant to be neither exhaustive nor particularly precise. Furthermore, some of these kinds of understanding are undoubtedly necessary (though not sufficient) for certain instances of explanatory understanding. Nevertheless, the devils in their details need not be exorcised here – my account of explanatory understanding should be compatible with whatever turns out to be the best accounts of these kinds of non-explanatory understanding.

My fixation on explanatory understanding includes one notable detour. Chapter 4 discusses explanatory understanding's relationship to something that philosophers call *objectual* understanding. Roughly stated, it is the understanding one has of a subject matter. It takes as its complement a noun phrase (e.g., "Niels understands quantum mechanics"). I will argue that explanatory understanding already captures anything philosophically important about objectual understanding. Hence, aside from considerations of linguistic convenience, we have no real need for the latter.

1.2. Degrees of Understanding

However, even when we focus on explanatory understanding of empirical phenomena, our philosophizing can only begin after a further clarification. Such understanding admits of *degrees*. For instance, recall our protagonist, Susan, who understands why the sky is blue. But now assume that she is a leading atmospheric physicist. Presumably, her understanding of the sky's blueness would be quite robust, involving a grasp of many causal factors, connections with deep theoretical principles, experimental results, methodologies, and so on. By contrast, we might credit Susan's freshman student, Bill, with understanding why the sky is blue even though he grasps only a tiny fraction of the information at Susan's disposal. In short, Susan's understanding is *better* than Bill's.

How do we navigate these different degrees of understanding? We might analyze a kind of *minimal* understanding by identifying the conditions that are necessary for any understanding whatsoever. Alternatively, we might analyze a *maximal* or *ideal* kind of understanding, which would be a mirror image of minimal understanding. In either case, a way of *comparing* different people's understanding would be a handsome prize, as we could then describe the full spectrum.

I will start with minimal and comparative principles of understanding, and then derive other gradations accordingly. The details of these principles needn't concern us quite yet. For now, I simply want to sketch how we could cover the gamut of understanding from my preferred starting point.

4 The Philosophy of Understanding

To begin, what does it mean to understand *better*? This is where comparative principles enter our happy scene. They will have the following form:

Schema for Comparative Understanding: Ceteris paribus, S_1 understands why p better than S_2 if and only if S_1 has minimal understanding of why $p + X$.

Obviously, this is just a sketch; we still need to fill in the value of X . Nevertheless, even these bare bones invite a few *obiter dicta*.

First, note that one can only have *better* understanding if one has *minimal* understanding. For instance, Susan could not understand why the sky is blue better than Bill if Susan does not have at least *some* understanding of why the sky is blue. This point seems obvious enough that I will leave it implicit hereafter.

Second, X may have several moving parts. Thus, some comparisons will be messy, as different individuals may have greater understanding with respect to one dimension of X , while being further away with respect to others. For instance, one person may have a tenuous grasp of the evidence that confirms her extremely accurate explanation; another may have a firm grasp on the evidence that supports her less accurate explanation. In this kind of situation, the *ceteris paribus* clause is violated, and there may be no clear way to determine who has better understanding. Alternatively, the relative importance of these differences may be specified by the context in which the comparison is made. Of course, in other cases, one understander strictly dominates another, and the comparison is relatively straightforward. This is presumably the case with Susan and Bill.

We can then use these comparative principles to derive a conception of ideal understanding:¹

Ideal Understanding: S ideally understands why p if and only if it is impossible for anyone to understand why p better than S .

Note that here, “impossible” means “logically impossible” and not merely “humanly possible.” This is an ideal, so we might as well aim high!

From this ideal, we can then pair a minimal account of understanding with a contextualist semantics to make sense of non-comparative or “outright” understanding, which earns its keep in the ample space between the minimum and the ideal:

¹ I’ve learned much about how to think about degrees of understanding from Kelp (2015). One small difference: Kelp takes maximum/ideal understanding as his starting point and gleans comparative and outright conceptions of understanding from there. These are only methodological points. I briefly cover the substantive differences between our views in Chapter 4.

1.2. *Degrees of Understanding*

5

Outright Understanding: “*S* understands why *p*” is true in context *C* if and only if *S* has minimal understanding and *S* approximates ideal understanding of why *p* closely enough in *C*.

Thus, on this view, contexts dictate how closely one must approximate the ideal. For instance, my understanding of why my car moves consists of little more than my facility in depressing the gas pedal. This understanding has proven serviceable on the boulevards of Vermont, but has not served me nearly as well in the arena of car repair. Thus, given the rules of the road, I understand why my car moves, but, given the standards in the shop, I do not. Thankfully, my local mechanic ably meets the standards in these latter contexts.

Furthermore, an account of outright understanding presupposes a prior theoretical account of *both* minimal *and* better understanding (the latter by way of ideal understanding.) If a context’s standards of approximating the ideal fall below the threshold of minimal understanding, then they are too weak to do their job. However, if outright understanding did not appeal to principles of better understanding, then it would be identical with minimal understanding.

Finally, I introduce another concept of understanding mostly for convenience. Sometimes, we need not specify how much understanding agents actually have. Suppose that a person knows her way around a car engine, but we do not know how deeply that understanding runs. In such cases, it will be useful to use “generic understanding” attributions:

Generic Understanding: *S* has some understanding of why *p* if and only if “*S* understands why *p*” is true in some context *C*.

In other words, we can think of generic understanding as having understanding to some degree or another.

As we will see, much of the extant philosophical literature on understanding frames its questions as if understanding were an all-or-nothing affair. Thus, the key questions have been, “Does understanding require explanations? Must it consist of mostly true information? Is it a species of knowledge?” I will be recasting many of these questions so as to capture the fact that understanding admits of degrees.

To summarize, I will take accounts of objectual, ideal, outright, and generic understanding to spring from the fountainhead of minimal and comparative accounts of explanatory understanding. Let’s begin to build this fountainhead, beginning with two comparative principles.

1.3. The Nexus Principle

A natural suggestion is that explanatory understanding is the possession or “grasp” of an explanation. For instance, to understand why the sky is blue is to have a correct explanation of why the sky is blue. However, multiple factors contribute to the azure above. For instance, blue light’s wavelength is relatively short, it is scattered in all directions by molecules in Earth’s atmosphere, the spectrum of light emission from the sun does not distribute all frequencies in equal proportion, the high atmosphere absorbs violet light, individuals’ positions to the sun changes at different times of the day – and don’t get me started on our eyeballs’ intimate workings.

Presumably, one’s understanding of the sky’s blueness increases as one gathers more of these correct explanatory factors and also as one learns how these correct factors hang together. Let the *explanatory nexus* of p be the set of correct explanations of p as well as the relations between those explanations.² I suggest the following as the first of my two comparative principles of understanding:

*The Nexus Principle: Ceteris paribus, if S_1 grasps p ’s explanatory nexus more completely than S_2 , then S_1 understands why p better than S_2 .*³

This raises three important questions. First, what is a correct explanation? Second, which relations between explanations furnish understanding? Third, what makes one person’s grasp more “complete” than another’s? I address each in turn.

1.3.1. A “Theory” of Explanation

I face an interesting challenge in depicting the nexus’ main inhabitants: correct explanations. On the one hand, because I would like the larger points about understanding to swing freely of any of my idiosyncrasies about explanation, this favors being relatively noncommittal about the nature of explanation. On the other hand, if I am too noncommittal, then my claims about understanding become inscrutable. In writing this book, the following seems to have struck the right balance:

² In conversation, it appears that some people individuate explanations slightly differently than I do. If you find yourself stumbling on this, then simply replace the count-noun “explanations” with the mass-noun “explanatory information.” Thus, the nexus of p is the totality of explanatory information about p , and grasping more of this information improves one’s understanding.

³ Recall from Section 1.2 that I am leaving implicit the requirement that S_1 has better understanding than S_2 only if S_1 has minimal understanding. Parallel points apply throughout my discussion of better understanding.

1.3. *The Nexus Principle*

7

q (correctly) explains why p if and only if:

- (1) p is (approximately) true;
- (2) q makes a difference to p ;
- (3) q satisfies your ontological requirements (so long as they are reasonable); and
- (4) q satisfies the appropriate local constraints.

Hereafter, “explains” is elliptical for “correctly explains,” unless otherwise noted. Furthermore, I will follow the time-honored philosophical pretension of using the Latin *explanandum* to denote p (the statement to be explained) and *explanans* to denote q (the statement which does the explaining).

Let’s discuss each condition in turn. The first condition is relatively uncontroversial. For instance, nothing correctly explains why the sky is paisley, why Mitt Romney won the 2012 presidential election, or why electrons are positively charged.

Similarly, many theorists of explanation agree upon the second condition. Consider the claim that blue light’s short wavelength explains why the sky is blue. This entails that blue light’s wavelength *makes a difference* to the sky’s color. A common way of unpacking this is in terms of counterfactual dependence: had blue light’s wavelength been longer, then the sky would have been a different color (Lewis 1986; Woodward 2003). For the most part, I will assume this counterfactual approach to difference-making. However, my arguments should not be affected if an alternative approach to difference-making (e.g., Strevens 2008) were countenanced in its stead.

The third condition, that the explanans satisfy your (presumably reasonable) ontological requirements, is designed to elide complicated issues concerning scientific realism. For the purposes of this book, realists will hold that the explanans q should be treated in the same manner as the explanandum – it should be (approximately) true. By contrast, many antirealists deny that our best explanations have true explanantia. A prominent antirealist alternative only requires them to be empirically adequate (van Fraassen 1980). Empirical adequacy is, roughly speaking, the requirement that a theory says only true things about directly observable entities, processes, and the like. This applies not only to those phenomena that are actually observed, but also to those that are *observable* – including all past, present, and future phenomena. By contrast, anything a theory says about *unobservable* entities – paradigmatic examples of which are subatomic particles, the curvature of spacetime, species,

8 The Philosophy of Understanding

mental states, and social structures – may be false without forfeiting explanatory correctness.⁴

Many philosophers would chafe at this proposal and require that both the explanans and the explanandum are approximately true. I will be walking a fine line here: none of my arguments obliges someone to abandon this stronger commitment if he or she is so inclined. However, neither do my arguments compel someone to adopt these stronger commitments if he or she is disinclined. Feel free to swap out this third condition with a requirement that satiates your realist longings without suspicion or shame. For all practical purposes, this will only bear on Chapter 6.

Finally, we get to the fourth and most cryptic of my requirements on explanations – that they satisfy “local constraints.” I take the first three conditions on explanation to be “global” constraints: they apply to explanations anywhere we find them. But, gaze deeply into my soul, and you will see a card-carrying explanatory pluralist staring back at you: the relevance of many explanatory features depends on the specific explanandum, the standards of the discipline, and the interests of the inquirer. To get a taste of these local constraints, in subsequent chapters we will see that only some of our explanations:

- represent causal structure;
- deploy asymptotic reasoning;
- represent mechanisms;
- represent non-causal, contrastive, probabilistic relations
- unify phenomena into a single framework;
- use idealizations; and
- represent potential interventions.

I emphasize that these local constraints must be satisfied *in addition* to the three global constraints I place on explanation. Hence, I take the global constraints to assuage concerns that I’m being too slippery, and the local constraints to afford me enough flexibility to remain faithful to the diversity of scientific practice and to grant certain assumptions to my interlocutors that will enliven the dialectic by avoiding foot-stamping impassés about what is – and is not – an explanation.

⁴ The labels “scientific realist” and “scientific antirealist” aren’t the tidiest tags to use here. For instance, there are entity realists (Cartwright 1983; Hacking 1982) and structural realists (Worrall 1989) who reject the idea that our best explanations are approximately true. These folks might well count as antirealists in this book.

1.3. The Nexus Principle

9

Figure 1.1

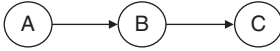


Figure 1.2

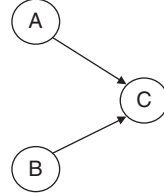


Figure 1.3

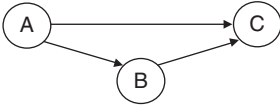
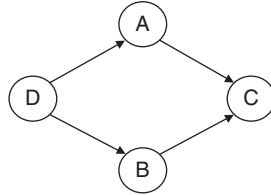


Figure 1.4



Figures 1.1 to 1.4: Different Explanatory Structures

A letter at the tail of an arrow does the explaining (explanans); a letter at the head of an arrow is explained (explanandum).

1.3.2. Inter-Explanatory Relationships

The explanatory nexus includes not only correct explanations, but also the *relationships* between them. What is characteristic of these relationships? One obvious relationship is that of *relative goodness*. Some explanations are *better* than others, even if both are correct. For instance, the presence of oxygen is explanatorily relevant to any fire's occurrence. However, very rarely will the presence of oxygen be the *best* explanation of a fire, in part because oxygen is also frequently present in the absence of fires. Per the Nexus Principle, grasping these sorts of facts enhances one's understanding.

However, superiority and inferiority are not the only relationships between correct explanations. For instance, consider the four scenarios in Figure 1.1. to 1.4. Suppose that in all of these cases, both A and B are correct explanations of C. However, even if, e.g., A is a better explanation of C than B, this would not say anything about the *structure* that these diagrams represent.⁵ Quite clearly a person who could not distinguish these different

⁵ Perhaps indirect explanations of C, such as A in Figure 1.1 or D in Figure 1.4, should not be part of the nexus. I submit that if this is so, then they will figure in one's scientific knowledge of the nexus. As sketched in Section 1.4 and discussed more extensively in subsequent chapters, this is also part of my account of understanding. Consequently, grasping indirect explanations will figure somewhere in my account of understanding; exactly where is not terribly important.

explanatory structures would not understand C as well as someone who did. For instance, a person who knew that A only explains C through B in Figure 1.1, or that A and B are independent of each other in Figure 1.2, or that D is a common explanation or “deep determinant” of both A and B in Figure 1.4, and so forth. Intuitively, this person has a better understanding of C than a person who did not grasp these relationships. Undoubtedly, explanations can stand in other relationships that figure in the nexus.

1.3.3. *Completeness of Grasp*

The Nexus Principle pins the goodness of one’s understanding to the *completeness* of one’s grasp of explanatorily relevant information. But what does the completeness entail? I submit that a person’s completeness of grasp is proportional to each of the following:

- The number of correct explanations and inter-explanatory relations grasped
- The quality/importance of the explanations and inter-explanatory relations grasped
- The level of detail of the explanations and inter-explanatory relations grasped

It would be a mistake to think of my account of completeness in quantitative terms – e.g., we do not typically count the number of explanations, nor will anything so detailed will be required. For the most part, we will encounter situations in which there is a stock of explanatory information that two or more inquirers both grasp and then some further bit of explanatory information that is unique to one. Indeed, science frequently progresses along just these lines. For the purposes of this book, these sorts of comparisons require no quantitative metric of one’s grasp of the nexus.

1.4. **The Scientific Knowledge Principle**

Essentially, I have taken a truism – that explanations are answers to explanation-seeking questions – as the springboard for the Nexus Principle. However, understanding involves more than this. The Nexus Principle appeals to the capacity of an agent to “grasp” explanatory information. However, to analyze understanding in terms of grasping is to swap out an enigma for a mystery.