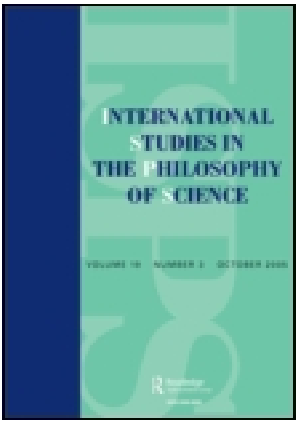

International Studies in the Philosophy of Science
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cisp20

An Inferential Model of Scientific Understanding

Mark Newman [a]
[a] Department of Philosophy , Rhodes College
Published online: 15 May 2012.

To cite this article: Mark Newman (2012) An Inferential Model of Scientific Understanding, International Studies in the Philosophy of Science, 26:1, 1-26, DOI: 10.1080/02698595.2012.653118

To link to this article: http://dx.doi.org/10.1080/02698595.2012.653118

International Studies in the Philosophy of Science
Vol. 26, No. 1, March 2012, pp. 1–26

Routledge
Taylor & Francis Group

An Inferential Model of Scientific Understanding

Mark Newman

In this article I argue that two current accounts of scientific understanding are incorrect and I propose an alternative theory. My new account draws on recent research in cognitive psychology which reveals the importance of making causal and logical inferences on the basis of incoming information. To understand a phenomenon we need to make particular kinds of inferences concerning the explanations we are given. Specifically, we come to understand a phenomenon scientifically by developing mental models that incorporate the correct causal and logical properties responsible for the causes or logical properties of the phenomenon being explained.

1. Introduction

Michael Friedman (1974) posed a challenge for any adequate theory of scientific explanation: show how explanations can provide us with scientific understanding. He argued that the deductive-nomological (D-N) account of scientific explanation advocated by Carl G. Hempel (1965) fails to explicate this relation because although D-N explanations may provide a sense of rational expectation they fail to generate understanding (flagpole heights can be predicted but not understood in terms of shadow lengths; barometer readings can be predicted but not understood in terms of incoming storms, etc.). Friedman advocated an alternative theory: the unification account of explanation. He argued that the psychological state of understanding is not a matter of derivation, but of reducing the total number of independent phenomena we accept as brute in our theories. The account he generated did not succeed but his argument highlights the importance of clearly explaining the connection between scientific explanation and scientific understanding.

Several new theories have been proposed to explain this connection. In Section 2, I discuss a couple of these accounts to illustrate the point that there is a mistaken

Mark Newman is at the Department of Philosophy, Rhodes College. Correspondence to: Department of Philosophy, Rhodes College, 2000 North Parkway, Memphis, TN 38112-1690, USA. E-mail: newmanm@rhodes.edu

ISSN 0269-8595 (print)/ISSN 1469-9281 (online)
http://dx.doi.org/10.1080/02698595.2012.653118

assumption in how to approach the issue. I argue that each account gives us a theory of how to *identify* when someone understands something, but they do not tell us what understanding actually *is*. That is, these accounts try to indicate how we can recognize when understanding has been achieved, but they do not provide a *constitutive* story telling us how scientific understanding is cognitively constructed from a scientific explanation. This is important, I argue, because we must appreciate the construction process if we are to account for the cognitive achievement of moving from an explanation of ϕ to an understanding of ϕ. In Section 3, I describe some important and relevant cognitive theory from which we can develop an inferential model of scientific understanding. I then unpack my inferential account of scientific understanding in Section 4. In Section 5, I illustrate how the inferential account works for some simple examples of good scientific explanations.[1]

2. Two Current Accounts

I select for analysis the accounts given by J. D. Trout (2002, 2005) and by Henk W. de Regt (2004, 2009; de Regt and Dieks 2005), because these two theories reflect opposite ends of a common spectrum where understanding is identified by what people can *actually do*—their abilities to answer questions, make predictions, etc. They are also the most developed of current accounts, and hence give us the most to work with. These accounts are also respectively objective and subjective in character—Trout takes understanding to be an objective property of the subject; de Regt rejects this view, advocating a pragmatic and contextually flexible theory. The selection of objective and subjective accounts provides us therefore with two quite distinct targets.

Trout's primary concern in his work is to show that the *sense* of understanding does not have epistemic worth. Since sometimes the psychological sense we get of understanding an explanation may be illusory, he argues, we should only be willing to lend epistemic credence to *genuine* understanding, that is, correct understanding. I won't spend time on his arguments for that particular thesis, but instead focus on his positive account. He offers us the following set of conditions for *genuine* understanding:

1. the belief putatively understood is at least approximately true,
2. the agent has sufficient collateral theoretical knowledge or information relevant to that belief, and
3. the belief is produced by a reliable process, perceptual or cognitive. (Trout 2005, 203)[2]

The first task in evaluating these criteria is to disambiguate the propositions. One can imagine the first proposition raising concerns for anyone suspicious of approximate truth, but the meaning is quite clear. Trout's commitment here is to *facticity*. He demands that if one is to understand some phenomenon *p* then one must have a set of beliefs about it that are true, or as true as we can make plausible. This contrasts with the claim that we can understand *p* even if we have an explanation for *p* that is *clearly* false. For Trout, then, Priestley failed to understand combustion because his phlogiston theory was not even approximately true, whereas Newton understood ballistics because, although strictly speaking false, his classical mechanics is approximately true. Although in general we would like to possess a complete theory of approximate truth, it would be inappropriate to reject Trout's characterization simply because we don't yet have such a theory. It at least seems plausible to think one fails to understand *p* if one's explanation for *p* is *mostly* false. After all, how can I seriously concede that you have achieved an understanding of classical electrodynamics if your beliefs about that theory include essential commitment to the luminiferous ether? We might accept that you have *partial* understanding, but just as with piecemeal approaches to scientific realism, we would be reluctant to say you understand since your knowledge claims hinge on clearly erroneous beliefs. So, it is best to accept Trout's first requirement on understanding, that we require beliefs be approximately true in regards to *p*, even though we recognize some flexibility in this demand.[3]

The second of Trout's requirements is less clear. He appeals to several important notions. The first is 'sufficiency'. He says the agent is to have sufficient knowledge or information relevant to their belief. But what does *this* mean? Is the requirement supposed to be that the agent has background knowledge enough to permit derivation of the relevant belief? Perhaps his concern here is with some foundationalist or coherentist notions of how to provide reasons that justify holding the belief. If so then it is unclear what the relevant notions are supposed to be, or what would amount to a sufficiency requirement on them. Then there is the inclusion of 'theoretical' in the clause. One is left to assume that this is in contrast to 'non-theoretical' along observable–unobservable lines, but nothing is said to defend this interpretation. Perhaps 'theoretical' is to be taken in a Lewisian sense as 'not yet defined in a prior theory', but who knows? So much is left unsaid about this requirement that it seems uncharitable to spend time attacking any one interpretation we give to it. Still, perhaps the idea generally ought to be that we should expect of someone that understands *p* not only that they have true beliefs about *p* but also that they have *supporting* beliefs about *p* (where 'supporting' is left open, but incorporates an epistemically important concern, be it foundationalist or coherentist, empiricist or realist). If that is the case, then it looks like Trout is on safer ground and we can accept (2). It would after all be peculiar to claim one understood *p* if all one knew was *that p*, rather than *why p*.

The last of his requirements is that the belief be produced by a reliable process, perceptual or cognitive. So, where requirement (2) seemed to reflect Trout's concern with internalist epistemic issues, (3) is distinctively externalist, appealing to (an as yet unarticulated form of) process reliabilism. This is not the place to debate the virtues and vices of internalist or externalist epistemology, but we are facing at least two serious concerns with regard to how this last requirement fills out his overall set of conditions. First, there are clear tensions between internalism and externalism, and if his second and third requirements are advocating on either side of this debate, then much more needs to be said to defend a synthesizing move like this.[4] Of course this is to assume a synthesis is intended, and perhaps Trout is able to respond in a way that avoids such a reading. Nevertheless, and this is my second concern, given what has so far (I think charitably) been said about requirements (1)–(3), it is unclear how what we have been given by him is anything but an attempt at providing a *theory of knowledge*, not an account of scientific understanding. His theory points to an agent having approximately true beliefs that are appropriately supported and derived by a reliable process. And this sounds an awful lot like an account of knowledge. And if that is his intention then something is surely amiss. Granting that understanding may be a form of knowledge, what makes anyone think it is *merely* knowledge? It seems clear that understanding is not identical to knowing since merely memorizing an explanation is very different from understanding it. Still perhaps knowledge of an explanation has a more complicated relationship to understanding it, and maybe the former somehow entails the latter.[5]

Does knowing an explanation for *p* entail understanding *p*? Take for example the phenomenon of a rainbow. Call the rainbow phenomenon, *p*. Now, we may possess many potential explanations for *p*, and perhaps one of them is even approximately true. Perhaps our explanation includes Snell's law of refraction, and combines this law with the appropriate law of reflection and the initial conditions of light rays passing through spherical water drops. If so, then our explanation can have a D-N form, with the conclusion being *p*. Assuming Snell's law is approximately true, and that we have good background beliefs supporting it, and that we formed the belief that *p* via a reliable process, then Trout's conditions are satisfied. That is to say, one can satisfy his requirements for understanding a rainbow merely by appeal to the old D-N model of explanation—appropriately combined with internalist and externalist constraints. Despite such constraints though, it is hard to see how this comes even close to providing us with understanding. The reliable process clause (3) reveals nothing regarding why we should believe *p* must follow from Snell's law plus initial conditions. So perhaps it is clause (2) that is doing all the work. But if that is the case, then merely knowing that *p* follows from the *explanans* is enough for understanding on Trout's account. That is, if all one additionally needs, given (3) is satisfied, is sufficient background beliefs to enable derivation of *p*, then one supposedly can understand *p* yet fail utterly to know why it is that Snell's law combined with those initial conditions must lead to *p*. One could fail to see *why* the cause is causing the effect, even though one might well know *that* the cause causes the effect. But the two are importantly different. To know that something causes an effect is to know *that p*, but to understand *why p* one needs something more than mere knowledge 'that'. Otherwise it would be plausible to think a physics novice who memorizes Snell's law, some initial conditions, and some consequence *p*, understands why *p*. He doesn't, he has merely memorized a series of facts. You may as well say he understands how to win a grandmaster chess match simply because he memorized a single game. Understanding *p* is clearly very different from merely knowing the explanation for *p*. Hence, possessing an approximately true explanation for *p* does not entail understanding *p*.

The upshot of this discussion is the following: Trout's criteria for *genuine* understanding are not sufficient for understanding. They capture our desire for a factive account that treats understanding as a form of knowledge, but they are deficient in that they fail to differentiate between knowing an explanation for *p* and understanding *p*. We require a theory that can fill this gap.

One might think that de Regt offers just such a theory. He certainly recognizes a similar deficiency in Trout's view. When considering Trout's claim that one can understand why a plane stays aloft along D-N lines, de Regt points out that 'merely knowing Bernoulli's principle and the background conditions does not suffice for explanation' (de Regt 2009, 26). But how are we to fill this gap between knowing an explanation for *p* and understanding *p*? De Regt and Dieks (2005) advocate a contextual account of scientific understanding. They claim that an agent understands *p* if he knows an intelligible theory of *p*. De Regt (2009) defines the intelligibility of a theory as the value that scientists attribute to the cluster of virtues (of a theory in one or more of its representations) that facilitates the use of the theory for the construction of models. Furthermore, the test for whether a theory is intelligible for scientists (in a context) is whether they can recognize qualitatively characteristic consequences of the theory without performing exact calculations (de Regt 2009, 33).

So how does de Regt fill the gap between having an explanation for *p* and understanding *p*? He says the gap is filled by constructing an explanation, and that '[t]he extra ingredient needed to construct the explanation is a *skill*: the ability to construct deductive arguments from the available knowledge' (de Regt 2009, 26). For de Regt (and Dieks) the difference between knowing an explanation for *p* and understanding *p* lies in one's ability to construct explanations for *p* using a theory that is intelligible to us.

Notice straight away the change in emphasis. Where I have been focusing on the difference between knowing an explanation and understanding its *explanandum*, de Regt locates the problem as being between knowing an explanation and *having the skills necessary to construct an explanation*. This is a very different issue, and I am not sure it is correct to place the emphasis on such problem-solving abilities. Going back to the airplane example, de Regt suggests that to understand why a plane stays aloft one must be able to use Bernoulli's principle and background conditions in the right way to derive the *explanandum* (de Regt 2009, 26). It is I think a mistake to be so demanding. To see this, imagine I have knowledge of both Bernoulli's principle and the initial conditions for a plane in the sky.[6] Imagine I also know that the plane staying aloft is a consequence of this principle being applied to these initial conditions. If we construct the explanation in a D-N format we can say that I see the conclusion follows from the premises: airflow over the top of a wing moves faster than that under the wing; faster moving air produces less pressure; therefore, airplanes stay aloft, since the pressure under their wings is higher than above their wings. However, it seems plausible that at the same time I am myself unable to actually *generate* any derivations using that theory—perhaps I am just incompetent at performing derivations, even though I can recognize how others achieve them. Surely, in this case where I have knowledge of the explanation but can produce no derivations using it, I understand why the plane stays aloft. If so, I have satisfied our conditions for understanding, although I have failed de Regt's conditions. I have knowledge of a theory, that theory is intelligible to me in the sense that I can see how it is used to make derivations, and I understand the explanation for the *explanandum*. I just cannot perform any derivation myself. I cannot *use* the theory to generate newly derived consequences for similar situations. I lack an important skill. However, it does not seem that I lack understanding of why the plane stays aloft.

If you are not yet convinced, perhaps unpacking this example as a scientific explanation will help illustrate why the problem-solving account is too demanding for scientific understanding. (This will also help to illustrate points that follow in Section 3.) Just how does a plane stay aloft? Here's an explanation from NASA's Glenn Research Center instructional website:

> Lift is the force that directly opposes the weight of an airplane and holds the airplane in the air. Lift is generated by every part of the airplane, but most of the lift on a normal airliner is generated by the wings. Lift is a mechanical aerodynamic force produced by the motion of the airplane through the air. (NASA 2010)

Moving on to explain how this lift is itself generated:

> Lift occurs when a moving flow of gas is turned by a solid object. The flow is turned in one direction, and the lift is generated in the opposite direction, according to Newton's Third Law of action and reaction. Because air is a gas and the molecules are free to move about, any solid surface can deflect a flow. For an aircraft wing, both the upper and lower surfaces contribute to the flow turning. (NASA 2010)

Furthermore, since the molecules of air are free to move around the wing and fuselage, forces between the aircraft and the fluid are transmitted back and forth at all points on the surface of the body. This means the overall transmission of force can be characterized as pressure. To determine the net mechanical force (**F**) on the aircraft, we integrate pressure (p) in the normal direction (**n**) for infinitely small sections over the entire surface area (a) of the body: $\mathbf{F} = \int p \, \mathbf{n} \, \mathrm{d}a$. Drag is the component of net force in the flow direction while lift is the component perpendicular to the flow. If the overall net force is in the upward direction, the aircraft stays aloft.[7]

In summary, for an aircraft passing through a fluid, like the air, there are mechanical forces acting at every point on its surface, which is best characterized as pressure. The net force can be calculated by integrating the pressure around the entire aircraft. If this net force is greater than the weight of the aircraft, it will stay aloft. In this explanation we are told how the mechanics of lift operate, especially the general idea of how one goes about calculating force values for how an aircraft stays aloft. There is, however, far too little information for anyone without an extensive background in theoretical physics (and a skill for adapting their highly developed abilities to aerodynamic problems) to actually perform a derivation for whether any given aircraft is going fast enough to stay aloft—a minimal problem task of the sort de Regt seems to have in mind. Given this explanation, it is plausible to think one can know an explanation for why planes stay aloft, yet fail to be capable of solving problems in aerodynamics. It seems that since we can follow this scientific explanation without actually being able to use the theory to solve further problems, the problem-solving account of understanding must be too strong.

One might worry that I am being too demanding in my interpretation of de Regt's position; after all he only calls for the ability to recognize qualitatively characteristic consequences of a theory without performing exact calculations. Deriving the net force on an aircraft by performing a complex integral is surely not what he has in mind. But if he is sticking to the idea of problem solving as reflecting understanding, I am not sure what kind of a scientific explanation he can possibly have in mind. Many theoretical explanations given in science are going to be complex—at least on a level with the integral involved in the calculation of lift stated above. If we minimize the level of understanding in any given case to be of a simple layperson's explanation of lift, then although one may be quite capable of recognizing qualitative consequences for a theory, it is not obvious that what one understands is a *scientific* theory at all. For instance, I might be able to 'see' that if one increases the velocity of an aircraft, all other things being equal, the lift force is going to increase linearly. This, however, is a popular-level relation, with little scientific explanatory import. Note that this explanation is not wrong; it is just not the sort of thing we are trying to capture with a theory of *scientific* understanding.[8]

One might worry then that my reading of de Regt is unbalanced because he claims only that having a theory with which to qualitatively determine the phenomenon is a *sufficient* condition for understanding. It seems instead that I am mistakenly taking the intelligibility of a theory as a necessary condition. But as previously pointed out, if we have only qualitative knowledge of why a plane stays aloft, then this is not much of a scientific explanation. So, although qualitatively determining the outcome of a theory is sufficient for a very weak notion of understanding, it cannot capture what we mean when attributing to someone the cognitive achievement of scientific understanding.

Here is my diagnosis of what has gone wrong in de Regt's account: in usual cases of scientific understanding we require that a subject *comprehend* the explanation given, but de Regt is demanding something much more of us—he requires that we also be able to *problem solve*. The two skill sets are significantly different, as I am about to show.

In summary, what the above discussion teaches us is that Trout's appeal to genuine understanding is insufficient for distinguishing between merely knowing an explanation for *p* and understanding *p*. On the other hand, de Regt's account of understanding as requiring problem-solving skills is far too demanding because it stipulates conditions that are not necessary. These theories therefore swing too far in either direction. I suggest that what we need is an account of understanding *p* which tells us not what it is to know an explanation for *p*, nor that we must be able to problem solve for *p*, but rather what it is to *comprehend p*. I turn to that task now.

3. Comprehension and Problem Solving

Assuming I am correct and what we require for a theory of scientific understanding is an account of how we comprehend a phenomenon, where might we start? What resources can we draw upon to construct our theory? Perhaps a good departure point is the case of someone learning from a textbook. This doesn't seem like a bad idea, since when learning a new scientific theory one is typically going through the process of coming to comprehend the world via theories and examples explained in a text. So, by taking as our starting point the situation of a student learning to comprehend some phenomenon *p* in virtue of reading about it in an expository work, we are using the most elementary case from which to build.

But before we begin thinking about what a student is doing when learning science, it is wise to ask if this process hasn't been studied before. You probably won't be surprised to hear that there is in fact already a well-established literature on the subject in cognitive psychology. (I am surprised that philosophers haven't made use of this work already. It seems eminently relevant.) In what follows I will describe some pertinent theory and new results that explain what has been learned about how students come to comprehend the world through science texts. I will then move on to drawing conclusions from this research and build a philosophical account of scientific understanding on that basis.

We can begin quite generally with some important distinctions made by cognitive scientists when discussing the process of acquiring understanding from a science-text explanation.[9] The first distinction is between *shallow* and *deep* knowledge.[10] *Shallow knowledge* is literal knowledge of an explanation. This may include the explicitly mentioned ideas in a text such as the definition of concepts, simple facts or properties of a concept or system, and even the major or large-scale steps in a procedure. One can develop shallow knowledge with mere referential inference on the information given. This level of comprehension is at the semantically atomic level, where concepts lack inter-theoretic integration and there is a minimum of coherence achieved. *Deep knowledge* is, however, less straightforwardly defined. We know that it requires activating cognitive processes that enable the learner to build one form or another of cognitive representation of the situation being explained. Deep knowledge is achieved by the encoding into memory of detailed coherent explanations. This knowledge once achieved permits the subject to perform further inferences, solve problems, make decisions, make predictions, etc. To appreciate even the rudiments of how these processes work and what these representations are, we must look at what has been discovered about their respective properties.

Beginning with cognitive *representations* of the explanation, these can be divided into *levels* and *kinds*. Along the levels dimension, there are: *surface code; explicit propositions*; *mental models*; *problem models;* and *pragmatic interaction*. I will begin by looking at the first three levels of representation because they are importantly different from the last two. *Surface code* is the least complex, and includes the syntax and wording of the text, and the lines, angles, shapes, texture, etc., of images. The *explicit propositional* level is next, and this refers to our success at capturing the meaning of text and images. This is often called the *textbase*. The third level is that of the *mental model*, also known as the *situation model*. This model is a mental representation of the explanation which includes among other things the causal sequence of events that unfold, the function of components in the sequence, and if there are intentional agents involved it will include their goals and purposes.

We can make this more concrete if we go back to our example of why a plane stays aloft. The surface code in the explanation is the story we are told in its most literal form—the words used, such as 'plane', 'Bernoulli's principle', etc. Surface code would also include any images incorporated, such as a plane, its wings, air flowing over the surface of the wing, etc. The textbase is the propositional content of the explanation—this is often represented by researchers in terms of predicate–argument pairs. For example, the proposition 'the plane flew through the air' can be represented as [FLEW THROUGH (Plane, Air)].[11] The third level of representation, the situation model, incorporates the causal, logical, and perhaps mathematical components of the explanation and their functional interaction. The plane is going through a process, flying, which is complex. The cause of its staying aloft intimately relies on the interaction between air flowing over and under its wings, and the related pressure differential that results from different flow speeds.

So far these levels of representation are constructed using relatively low-level cognitive processes, but there are deeper levels of representation. In particular, when it comes to scientific problems some researchers have identified a higher-level process where we construct what they call the *problem model*. This level uses not just everyday background knowledge, which is how the situation model is built, but also specific formal and scientific knowledge, such as is required to incorporate relations and values for variables in the statement of a problem. The idea here is that there is a difference between the subject building a situation model for some scientific explanation, and building a scientifically informed model that inherently incorporates scientific concepts and principles for the purpose of solving a particular problem. Going back to our example, a situation model can be built using causal and formal principles that describe how the airflow both under and over the wings creates more net lift force than is countered by weight or drag. A problem model builds a new model using the same set of concepts and principles (and perhaps additional ones as necessary) to apply in a *new* case—such as when constructing a model of how fast a larger plane has to travel to stay aloft (given its particular mass, wing surface area, shape, etc.). Importantly, the cognitive processes required for building a situation model are not identical to those required for problem solving in the sense required by de Regt's theory of understanding. Therefore we should not identify the building of situation models with that of solving problems, but more of this shortly.[12]

As I mentioned above, deep knowledge is not merely a matter of constructing a particular kind of *representation* at a particular level (as just illustrated), it is also importantly revealed via the types of cognitive *processes* involved in understanding an explanation more broadly. Although there is no uncontroversial matrix that establishes which types of cognitive processing go into building a type or level of representation, or what we do with them once they are built, it is established in the cognitive psychology literature that there are *simple*, *intermediate*, and *difficult* cognitive processes, and each is correlated with parts of model building and 'use activity' to some degree or another. I will briefly sketch the taxonomy of these cognitive processes (which are used for constructing the different kinds of representations just sketched).[13]

The most simple cognitive processes include *referential inferences* (recognition), and *recall*. For our example these are used when we identify in our model what a plane is, what Bernoulli's principle is, etc. The next, intermediate, type of processing is at the level of *comprehension*. This includes the generation of inferences necessary for building a model, be it a situation or a problem model, as well as the integration of the information coming in from the explanation into the broader background network of knowledge possessed by the subject. This type of processing is my primary interest in this article, and I will say more about it in just a moment. For now, however, it is important to recognize how comprehension is comprised of processes different in kind from those used in higher-level cognitive processes we see in the taxonomy. These even more demanding functions are at the highest level, and include: *application*; *analysis*; *synthesis*; and *evaluation*. I will say a little about each before moving back to unpack what has been learned about comprehension, which I take to be the appropriate source for a philosophical theory of scientific understanding. The essential points in what follows are that there is a significant difference between the intermediate- and high-level types of processing, and that comprehension processes, which are intermediate, are not to be confused with problem-solving processes, which are high level.

Attempting to sketch all that has been said about the higher-level cognitive processes would be ridiculous, but we can understand each of the four subcategories listed above better with reference to our plane example. *Application* is the process of applying knowledge derived from an explanation to a new situation, such as trying to solve an aerodynamics problem for a biplane instead of a conventional winged aircraft. *Analysis* is the process of breaking down the elements in an explanation and identifying the relations between them. For example, recognizing for our plane example how the pressure differences above and below the wings interact in important ways to enable the craft to stay airborne. *Synthesis* is the process of putting together new patterns or constructing new structures in providing a novel solution to a problem. Here the idea reflects how a student may go about solving in new ways a problem of calculating the necessary lift required. Lastly, *evaluation* is the process of judging the accuracy, effectiveness, or value of a solution to a given problem. This is done by appeal to some (usually subjective) criteria and standards with which the checking is performed. Here the student may reflect on the result he has calculated as a solution to the problem, and then evaluate whether it is a reasonable number given background assumptions.[14]

Each of these processes can of course be more or less demanding for the student, and consequently rank-ordering them in a strict manner would be controversial.[15] However, it is a popular view amongst those in the discourse and comprehension field to consider each of these higher-level processes as importantly different from the intermediate-level process of comprehension. This is important for my purposes because it reflects the difference between my account and de Regt's. In order to appreciate the differences between cognitive processes used for comprehension and those used for problem solving I now need to unpack some of the general characteristics of comprehension.

Cognitive psychologists differ over the details of comprehension, but their research focus has been primarily on the role of memory, inference, and coherence during the construction of mental models. For instance, they have investigated the properties and distinctions between the familiar categories of short-term memory, working memory, and long-term memory, as they relate to reading comprehension. There is now the new category of long-term working memory advocated by some on the grounds that it accommodates better the empirical results (Ericson and Kintsch 1995; Kintsch 1998). On this account the contents of short-term memory may at any time trigger processing in long-term memory to rapidly fetch additional content. For example, an expert who appears to have a particularly skilled working memory, may better be considered to have developed a skill for high-speed retrieval from the field with which they are familiar.

Not only has memory been of interest, but so too are the multiple kinds of *inferences* we make while working through science texts (Graesser, Singer, and Trabasso 1994; Cote, Goldman, and Saul 1998; Graesser and Bertus 1998). There are many classes of inference that have been identified, including: *anaphoric references*; *bridging inferences*; *explanation-based inferences*; *predictive inferences*; *goal inferences*; *elaborative inferences*; and *process inferences*. I won't go into details, but the most recent literature indicates that the causal and logical inferences being made while we process explanatory information play an important role in constructing coherent representations. This has led researchers to look carefully at how explanations might sometimes helpfully omit inferential steps so that knowledgeable readers have to make the coherence-increasing inferences themselves while building their mental models. On the other hand, sometimes it is inappropriate to leave such omissions, and more explicit step-by-step explanatory work needs to be provided for the student (McNamara et al. 1996).

Other areas of investigation related to coherence and inference are the role of images, models, and simulations in comprehension as well as the role of metacognitive processes that guide and monitor the model building itself, especially processes that determine what level of coherence is satisfactory for the agent. What is clear in the literature is that there exists a significant disparity between the kinds of cognitive processing involved in constructing a mental model, and those required for solving problems. Walter Kintsch, for example, a leading researcher in the field, says:

> I have insisted on a distinction between problem solving and comprehension. Comprehension is automatic, bottom-up, described by the mechanisms of the CI [construction integration] theory. Problem solving is a controlled, resource demanding process involving the construction of problem spaces and specialized search strategies. (Kintsch 1998, 394)

He further argues for this distinction on the grounds that although comprehension and problem solving both involve a construction phase and a solution phase, the nature of each phase is very different:

> For comprehension, the CI theory claims, the construction phase is essentially guided by the textual (or other perceptual) input. Propositions are constructed more or less closely, mirroring the input sentences of a text sentence by sentence. This is a highly constrained process, at least for the ideal reader ... The construction operations themselves are typically highly practiced and demand few mental resources. This is the realm of long-term working memory ... A very different

> situation obtains for true problem-solving tasks. Typically, the input vastly underconstrains the construction of problem representation. Instead of a text that almost dictates the kind of mental representation that will be constructed, the problem statement itself gives few hints as to what the problem space looks like . . . Often, the operators that must be used for the construction of the problem space are unfamiliar. The problem solver in this situation cannot rely on a few well-practiced, highly overlearned, automatic operations, and hence not on available retrieval structures in long-term working memory. (Kintsch 1998, 395)

The point should be clear: the resources of comprehension are not only less demanding than are those used in problem solving—reflecting my concern with de Regt's account—but they are of an entirely different kind. In comprehension we are using low-level processes such as integration and inference. This can be thought of as a relatively straightforward spreading activation process. Problem solving requires spreading activation too, but goes far beyond these relatively simple processes, also demanding of the cognizer much more directed processes, such as means–end analysis (Newell and Simon 1972; Newell 1990).

Kintsch is a leading figure in the field, and although what he says is not unchallenged his view is representative of a large part of the literature. Furthermore, where there are significant ambiguities between comprehension and problem solving, it is perhaps within the resources of those like Kintsch who defend this distinction to provide plausible explanations. For instance, when it comes to experts, it is sometimes hard to distinguish whether what they are doing is comprehending a situation at a very deep level, or solving a problem. As just mentioned, a popular example is of the chess master, who can simply look at a board configuration and see what move to play next. It is hypothesized that in this case the expert is relying on long-term working memory to such an extent that his 'solution' is derived automatically in a way more similar to the way most of us comprehend unchallenging explanations than to how we might solve a problem.

Similar solutions are proposed in accounting for decision-making errors. Kintsch uses the common example of Linda the bank teller to illustrate how this works for explaining the error of *representativeness bias*. We are given a description of Linda as an outspoken, intelligent 31-year-old, who is single, majored in philosophy, and was involved in college with a number of social-issue groups. We are then asked to judge which is more likely, that she is a bank teller or that she is a bank teller and is active in the feminist movement. Of course, many people mistakenly opt for the latter choice, although by the multiplication rule for probabilities of conjuncts we can easily see that it is more likely to be the former. The advocate of long-term working memory theory argues that the problem responses given by subjects match up with the way the story about Linda is cognitively processed. If a subject represents the description as a situation model, and if their decision is based on simple activation values in their cognitive representation, then the biased result is given. If the subject represents the story as a problem model, with formal high-level representation of the components of the story and appropriate elementary probabilities assigned, then they will give the correct solution. The upshot: studies in comprehension and problem solving can provide a theory as to why subjects give the responses they do to this problem. In particular, many subjects make the error of thinking it more likely Linda is both a bank teller and in the feminist movement precisely because they don't treat it like a problem; they treat it like a story to be comprehended.

This is just one example, and continued research in discourse and comprehension will help to confirm the difference, but it seems reasonable to take the above to show that comprehension can be understood as requiring a different set of cognitive processes than does problem solving. One method of verifying this last claim is to look at what distinguishes some learners who develop comprehension of a story from those who achieve either literal knowledge or problem-solving ability.

The results of experiments indicate that the most important difference between comprehension and either mere knowledge acquisition or developing problem solutions has to do with the inferences we make when constructing a situation model. To grasp this distinction a little better, consider the different types of inferences we make when coming to comprehend even a simple piece of text. There are several distinct kinds of inferences we make when coming to comprehend.

The kinds of inferences made when encoding information into long-term memory are recognition or recall inferences. They enable subjects to recognize what is being said or explained. These we can call 'referential inferences'. Referential inferences require some minimal level of comprehension, since without even the slightest understanding of any clauses in an explanation it is hard to see how our subject could even be capable of knowing what the explanation says. However, that's all we can claim for the subject who knows an explanation but fails to understand it. This would be the case if I had no knowledge of physics, but were only told that the reason the plane stays in the sky is 'because of Bernoulli's principle'. I really wouldn't have a clue what was going on, but I could memorize the explanation. I may know the explanation for *p* but utterly fail to understand *p*.

There is of course the other end of our spectrum, as represented in de Regt's criteria for understanding, where I am capable of making qualitatively accurate predictions about what will happen to the plane if Bernoulli's principle were suddenly to fail, or if the wing surface area was halved, or if the tail fell off. This seems to indicate understanding, but as I am sure by now is clear, this is not what constitutes understanding, even if it provides a means for identifying when a subject has it. Not only is the notion of understanding used in cognitive psychology far more constrained than this set of processes permits (as reflected in the quote from Kintsch) but this is also true of our everyday notion of understanding. We have already seen that the cognitive processes being used when we solve problems, such as application, analysis, synthesis, and evaluation, are of a much more demanding kind than are those implemented in the less strenuous task of understanding. The former processes work on problem models, and are far more controlled and deliberate. The processes that constitute comprehension are less demanding, quite often implicit to the learner, and may even be considered 'automatic'. Referential inference is one particularly implicit type of these processes, but there are others. Everyday examples of these implicit processes are easy to generate. Here are a few:

(a) The driver carelessly tossed his cigarette out of the window. The fire destroyed many acres of forest.
(b) Bob wanted to play music on his long train ride. If only he had charged his iPod before leaving the house!

These simple examples illustrate two other types of inferences: how implicitly we make both *causal* (a) and *logical* (b) inferences all the time. Everyday comprehension using these kinds of inference processes is a staple of our cognitive activity. Sometimes we come across an inferential step that requires just a little conscious thought, in which case the activity usually becomes explicit because it is more cognitively demanding. Importantly though, both implicit and explicit inference as an element in developing comprehension is very different from everyday problem solving.

One way to verify these differences is to look at what happens to an individual's comprehension of an explanation when they fail to make these particular types of inference (causal and logical). In one experiment for example it was found that the relation between comprehension and inference making was clear (Cain et al. 2001): children with comprehension difficulties are deficient at making inferences that require integration of text information with prior background knowledge. The less-skilled comprehenders' difficulties are not restricted to reading alone, but extend to auditory tasks. Neither is their poor comprehension explainable on grounds that their memory base was inferior to good comprehenders. It seems that the source of inference difficulties lays in their inability to select the relevant information for making correct causal or logical inferences—even though that information may have been in their memory store.

So where does all this get us? Thus far, I have argued that existing philosophical accounts of scientific understanding are inadequate at explaining the gap between knowledge and understanding either because they demand too little or too much. I have suggested that cognitive theories of comprehension might help us see this difference more clearly, and have provided some details on how we process information for comprehension according to current cognitive psychology. However, we have yet to see any positive philosophical account of understanding. In what follows I take what has been illustrated by the comprehension researchers and use it to develop a positive philosophical account for scientific understanding.

4. An Inferential Model of Scientific Understanding

As described above, cognitive psychologists distinguish between comprehension and problem solving by highlighting the different cognitive processes involved in each of these achievements. It is the hallmark of comprehension to build situation models by encoding information as mental models and using inferential processes that constrain as well as illuminate those models to varying degrees of coherence. But a scientific account of comprehension is not a philosophical theory of scientific understanding. What we need is a philosophical account that goes beyond the theorizing we have seen from cognitive psychology, one which provides an explanation for the established difference between knowing an explanation for *p*, understanding an explanation for *p*, and constructing an explanation for *p*. We have seen the first of these offered by Trout, and the third by de Regt. They both thought that they were providing the second. The pieces are in place now for a new account that gets the job done right.

I want to draw on what we have learned from cognitive psychology, which is that understanding a phenomenon relies on our making causal and logical inferences on the explanation of that phenomenon. Accordingly, the philosophical account of scientific understanding I advocate has the following underlying idea: *understanding a scientific explanation is a cognitive achievement constituted by our having made appropriate referential, causal, logical, as well as coherence inferences on information we encode into memory as knowledge.* That is to say, when we attribute to someone the state of understanding a scientific explanation for some phenomenon, *p*, we are not taking them to have merely literal knowledge of what is going on that generates *p*, nor do we think they must be able to derive *p* themselves from only initial conditions and laws. Rather, we are attributing to them the mental state of having made appropriate inferences on each part of the explanatory story that ends with *p* as its conclusion. We think they 'get' the meaning of each step in the explanation, and they have connected those steps with the right causal or logical inferences. We also think they have a coherent network of beliefs that now includes the explanation of *p*, rather than permitting them to make the correct inferences given in an explanation yet still have radically incompatible background beliefs.[16]

But what is it that makes the inferences the correct inferences? What constrains an agent in understanding the explanation given, rather than completely *mis*understanding it?

I suggest that what makes the agent's inferences the correct and relevant inferences for understanding *p*, be they causal or logical, is that *they correctly reveal the* reasons *that the causal and logical properties in the explanation are the causal or logical properties that they are*. That is, the agent's inferences must reveal to the agent which properties are *making* the causes of *p* the causes of *p*, or the logical properties of *p* the logical properties that they are. To put it more precisely, the correct inferences are those that infer to the properties that themselves explain why the causes or logical properties do the causing or entailing they do to generate *p*. So if an explanation says that in general *C*s cause *P*s, then to understand any given *p* requires that we infer the reasons why *C*s cause *P*s—to know what properties of any given *c* makes it such that it causes *p*. Similarly for logical properties *L* that lead to *p* having the logical properties that it does. This is what it is to make the correct inference regarding which properties make it causally/logically efficacious for *p*.

If this simple theory of understanding is correct, then it will explain the difference between knowing an explanation for *p*, understanding an explanation for *p*, and being able to generate an explanation for *p*. I think it does just this. Where an account like Trout's provides us with a theory of what it takes for someone to know how *p* is generated, my inference account of understanding says his story falls short of understanding. To understand *p* one not only needs to reliably form beliefs about *p* but also make the relevant inferences regarding its causes. This explains how a reliabilist reductive account of understanding is inadequate in accommodating the difference between knowing and understanding. On the other hand, my account does not require the agent be capable themselves of generating predictions about a phenomenon in order to understand it. We don't think someone must be capable of exercising the skills required to apply old knowledge about *p* to a new situation in order for them to understand *p*. This demands too much. It is enough that the agent be able to correctly identify the properties of the causes of *p* and identify how those causes work together to cause *p* in the given situation. We don't require the agent be capable of applying these causes to new environments. This contrasts with the problem-solving account, where in some situations one does not even require knowledge of the causes of the causes of *p* to generate predictions in a new situation. This seems too generous because in that case the agent may simply be problem solving in a 'plug and chug' manner rather than genuinely understanding. If such a mechanical procedure is permitted by a theory of understanding, so much the worse for that theory. My account of understanding as inferential can account for this distinction between understanding and mere computation.

Still, my sketch of what is required for scientific understanding is somewhat abstract. In the remainder of this section I will illustrate how my theory is applied to our example of how a plane stays aloft, and in the next section I provide some more examples of its application to simple scientific explanations. So, moving back to our example of wing lift, we can see how understanding the explanation requires an agent not only acquire knowledge that *p* occurs as a result of Bernoulli's equation plus initial and boundary conditions, we also see specifically which inferences (causal and logical) are made by the learner when they come to understand the explanation. We should also appreciate this process as being significantly different from those involved in problem solving for aerodynamic lift situations. Let me describe briefly what my theory suggests is going on in an agent's mind when he comes to understand the explanation given above.

First of all, the mental object being built is a mental representation of the explanation in the form of a situation model. As explained above, this is the third level of cognitive representation, above surface code and explicit propositions, but below problem models and pragmatic interaction. The construction of this situation model requires the implementation of low-level cognitive processes. The lowest level of cognitive process is referential inference and recall. As previously mentioned, for our explanation this includes inference about the meanings of terms like 'aircraft', 'fluid', 'air', 'gas', 'mechanical force', 'surface', 'pressure', 'lift', 'weight', 'wing', 'molecule', 'turning', 'integrate', etc. When an agent builds a mental model of the situation he is drawing on his own knowledge of these terms and embedding them into the model. Presumably there is significant overlap in what people connect with each of these terms, but none of our semantic networks is identical, and consequently no individual's mental model will be identical to another's.

The next level of cognitive processing is where we generate causal and logical inferences necessary for building our situation model. These inferences are generally automatic, and quite often go unnoticed by the learner. They are of a kind with those in our previous examples of (a) and (b). In our current aeronautical example the learner may generate an image of an aircraft, or even of a particular wing, turning air downward as it flies. If one is accustomed to force diagrams in physics, the image may be associated with arrows representing the force of lift generated by air pushing back on the aircraft, as well as of weight generated by gravity's influence on the massive object. The causal inference that it is the air *pushing* on the aircraft which gives it lift is an important step in comprehending the explanation for lift. In appreciating this inference the learner most likely is drawing on common-sense background beliefs about the relationship between notions such as 'push' and 'force' and 'motion'. For a more sophisticated student, the implications of lift may also involve an appreciation of how molecules impart momentum on surfaces, how molecules are not really perfectly elastic bodies with zero volume and point located mass, etc. But whatever level or depth of comprehension comes with the causal inference, one simply isn't grasping the mechanisms involved in the explanation if the situation model fails to include some causal properties that are responsible for the interaction between air and aircraft.

In more detail, the student who takes the explanation to be saying that lift comes from the air pushing back on the plane is making inferences when incorporating this information into a situation model. The inferences are positing the properties which are responsible for the air doing this pushing. In some cases of learning, which are quite shallow, the student merely appreciates that air pushing on a wing is going to impart a force, and if that force is large enough to overcome contrary forces, this will result in motion upward. The student who draws on more sophisticated background information relevant to the interaction between air and wing may infer that the molecules doing the pushing are themselves part of a Newtonian fluid stream which has mass, energy, and momentum. This second student has a deeper understanding of the explanation, but the important point for my purposes is that without any idea of the properties that cause air to push a wing upward, the student will have very little understanding of the explanation being given—they may *know* the explanation, but they will not understand it.

A similar point can be made for the part of the situation model which refers to calculating the net mechanical force on the aircraft. If one has no idea what an integral is, then one will fail to make a correct referential inference regarding calculating net force for one's model. One may in this case fail even to know the explanation in anything but a literal sense. On the other hand, knowing what integrals are allows many of us to know what it means to integrate pressure vectors around a surface. We are summing values for mechanical forces perpendicular to the body of the aircraft for infinitely small surface regions. To understand the explanation of how net force is calculated then, the depth of understanding is reflected by the number and importance of inferences we can make regarding the properties involved in the process of taking integrals for vectors. We need to know what rules apply for integration, what pressure means, what a vector is, etc. Just as with causal inferences, the more relevant inferences we make about the properties of logical processes, the more detailed and deep is our understanding of the phenomenon. The more detailed is our situation model too.

This doesn't mean, however, that we must be able to actually sit down and perform the necessary calculations to solve a problem—as seems to be the demand from de Regt. For instance, even though I am confident I know what it means to perform the calculation which sums the change in momentum and energy in a region of the fluid flow, don't ask me to actually do it. I couldn't. That doesn't mean I fail to understand the explanation for why a plane stays aloft. I *may* lack some depth to my understanding that someone who can perform this operation possesses, but it is absurd to think understanding totally evades me.

In summary, according to my inference model, to understand the explanation of why a plane stays aloft requires that, not only do I correctly make referential inferences about the textbase of the explanation, when building a situation model I also have to make inferences about the properties of the causal and logical *explanans* in the explanation that generate the *explanandum*. In our example these properties include the reasons why molecules impart momentum on impact, why pressure differences generate forces, what integration involves, etc.

This is a great deal more than is required for merely knowing the explanation, yet also less than a problem-solving account needs. On the latter, the learner not only builds a situation model, they are also capable of constructing a problem model. We have seen this difference illustrated for our example already, but in case the point needs reiterating: a problem model is significantly more cognitively demanding since its construction and solution requires the execution of cognitive processes such as application to new situations, analysis of elements in an explanation, synthesis of new structures in solving a problem, and evaluation of results obtained. For our example the task of solving such a problem is highly demanding, and certainly far more than one should require when attributing scientific understanding. Furthermore, I believe my account of understanding applies generally to all scientific explanations, and to support this boast I will next show how it might also apply to other examples from science.

5. More Examples of the Inferential Model at Work

The example of lift for an aircraft comes from elementary aerodynamics. Here are some more examples that illustrate not only how the inferential model works in different domains of science, but also how my account explains our diverging opinions over whether some explanations provide us with understanding at all.

We can start with our earlier example of explaining a rainbow. Here there are two things to explain: (i) how do we understand the phenomenon of white light splitting into a spectrum when passing through a water drop? (ii) What explains the bow shape of the rainbow? The answer to (i): white light is refracted when it first enters the drop. The angle of refraction, n, in general depends on the medium into which light enters as given by Snell's law, but it also depends on the light's wavelength for any given medium. The greater the light's wavelength, the lower is the angle of refraction. This fact results in a dispersion of the light as it passes from air into water. The split light travels through the drop, reflects off the back, and emerges at a point towards the front of the drop, again refracting upon its exit. It is the different wavelengths of light and the fact that they refract at different angles when entering or exiting a medium that explains the dispersion of white light into a spectrum. The answer to (ii): that we see an arcing rainbow, instead of some other shape, is a consequence of the fact (discovered by Descartes, no less) that rainbows form where we observe very high concentrations of light reflected at its maximum angle from raindrops. Because of the spherical geometry of water drops, the maximum angle at which light rays emerge has a very high ray density, much more than at other angles. We require this high concentration of rays in order to see the dispersion. For light passing through water drops this angle is 42°. That means we can only see dispersed light through raindrops when viewing from that angle. As we look out from our viewing location, points of dispersed light at 42° form a semicircle bow on the horizon: a rainbow.

What does the inferential account of understanding say about how we understand this example? Since the central idea is that understanding is achieved by making causal and logical inferences on the information given, what might these inferences be? First, assume we make the correct referential inferences about concepts involved in the explanation (we have pretty much the right ideas about what a spherical raindrop is, how refraction works, what a wavelength is, etc.). This gives us literal level knowledge of what's going on in the explanation. Now we build a (mental) situation model in which we causally and logically relate the components of the explanation by making a few important inferences: we first make an inference about the transition between air and water. The ray is bent and split in this process. These are causal notions. We know this causal process follows Snell's law, but if we don't know what this law says, then we significantly lack understanding about *why* the light refracts. If we know that the law says $n_{\mathrm{air}} \sin \theta_1 = n_{\mathrm{water}} \sin \theta_2$ then we are making progress, so long as we also know that n is an empirically determined value, θ stands for angles of incidence and refraction, and sin is a trigonometric function. But if we don't know what trigonometric function sin represents then we still haven't got very far in understanding the phenomenon. Even if we do understand the geometry, then there is still the causal question: why does light obey Snell's law this way? We are not given an explanation that answers this question, but for those who are able to make correct inferences about what it is that causes light to change angle and split while entering a drop of water, then surely we would consider them to have much deeper understanding than those who merely know *that* light does this. Making this substantial causal inference is inferring to the causes of why light in general is bent and split when changing media—inferring to the causes of the causes of the phenomena. This is what the inferential account requires of us if we are to have causal understanding of the explanation rather than mere knowledge. Most of us have something like a low-level logical understanding of the phenomenon since we are not sure why light is caused to refract.

The same goes for the next step in the explanation, which describes light's reflection at the back of the raindrop. Why does light reflect at the back of the drop? Those able to make further correct logical and causal inferences about this phenomenon achieve greater understanding. For instance, if one infers that the reflection involved in this case is determined by the angle of incidence and the index of refraction for water, then they will surely be said to logically understand rainbows better than those who have no idea how reflection works. And a similar story holds for part (ii) of the explanation. If one can make the logical inference that the angular radius of a rainbow is given by determining the angle of deviation for water droplets from Snell's law, then you will have a great deal more understanding of why the angular radius of the bow is 42°.[17] Thus, it is no mean feat to understand even a simple explanation like that given above. Importantly though, there is no requirement that we be able to actually derive correct predictions about the refraction or dispersion of light, or of why the radial angle of the rainbow is what it is. That would be to demand too much.

Now let's move on to an example from modern physics. Explanations in general and special relativity, as well as quantum mechanics have often been accused of revealing that we do not have genuine understanding of the world. The inferential account explains these differences. Take for example the gravitational redshift of light. Imagine a beam of light shone upward from the floor of an upwardly accelerating elevator in the gravitational field of the earth. We find that the frequency of light received at the top of the elevator is longer (redshifted) than its frequency at the source on the floor. Similarly, if the beam is shone down from top to bottom the shift is towards the blue. Why is this? Our best gravitational theories, and Einstein's general relativity in particular, can explain this phenomenon in a number of ways. One explanation is by appeal to Einstein's principle of equivalence, which claims that all effects of a uniform gravitational field are identical to the effects of a uniform acceleration of the coordinate system. Since light is subject to a relativistic Doppler effect (redshift) when source and receiver are in relative motion, the equivalence principle implies gravity will have the same effect. This entails that just as light suffers a redshift when source and receiver are in relative motion, light will also undergo redshift as it attempts to climb out of a 'gravitational well'.

Looking at this explanation for gravitational redshift, how deep is our understanding of the phenomenon? The inferential model suggests that the depth of understanding depends upon how many and how important are the inferences we make on the cause of the effect. Here we start with a description that requires the agent make referential inferences regarding concepts such as 'a beam of light', 'the gravitational field of the earth', 'uniform acceleration', 'coordinate system', 'Doppler effect', 'redshift', etc. The learner is also required to make causal and logical inferences on the relationship between accelerations and gravitational effects on light. This is where different levels of background knowledge play a vital role. If one is incapable of identifying the Doppler effect, as given by Newtonian mechanics, then there is little hope of acquiring much insight for the gravitational version. The classical theory says that if the source frequency of light ν is small compared with the speed of light c, then to first order in ν/c the observed frequency of light at the receiver ν' is $\nu' = \nu\ (1 + \nu/c) = \nu\ (1 + |\mathbf{g}|\ h/c)$, where $\mathbf{g}$ is the gravitational force and h is the height of the elevator. If a learner fails to draw on these exact relations, then although he may still get the general gist of redshift, we would say he lacks understanding in contrast to those who are familiar with these equations. Importantly though, for the novice who fails to make these connections an important consequence of the equivalence is going to evade understanding. This consequence is that the frequency of light waves emitted from a source depends upon the gravitational potential in which the light is propagated. Such a consequence can be appreciated, however, only if one recognizes that we can express the above equation in terms of the change in gravitational potential Φ between top and bottom of the elevator. Since $\mathbf{g} = -\text{grad}\,\Phi$ we can say $|\mathbf{g}| = -\Delta\,\Phi/h$. By substitution this gives us $\nu' = \nu\ (1 - \Delta\,\Phi/c^2)$, which in the Newtonian limit gives the gravitational redshift: $z = (\nu - \nu')/\nu$, and this shows $z_g = \Delta\,\Phi/c^2$. This is just to express that frequency depends on gravitational potential. None of which can be appreciated without making all the appropriate logical inferences regarding the variables in these relations.

Notice that these inferences all seem to be logical in nature, rather than causal. What constitutes a causal explanation for gravitational effects is a controversial issue. In the general theory of relativity we find explanations of particle motion not in terms of causal processes, but rather by appeal to the affine and metrical structure of spacetime and the variational equations of motion given by the theory—which are in turn explained in terms of the stress-energy tensor that describes the distributions of mass and energy. It is far from obvious that understanding a phenomenon should be limited to knowledge of only causal structure (*pace* Cushing 1991 and Salmon 1998) even if we would ideally like a clearer picture of how logic and reality merge in our best theories.

Take a final example to illustrate how understanding works, this time from outside of physics: how can we explain the rotation of bacterial flagella? The flagellum is the organelle for motility in bacteria, and often looks like a tail sticking out of the cell. The flagellum consists of three primary structures: a basal body (motor), a hook (universal joint), and a filament (propeller). The filament and hook rotate because they are driven by the rotation generated in the basal body. The body has an upper and a lower section. Two significant lower subparts that are responsible for generating rotation are the rotor and the stator. The rotor is attached to the upper subparts of the basal body, and through them to the hook and filament. The stator is connected to the cell membrane and remains stationary. Rotation is generated by either proton or sodium motive force generated between the stator and the rotor. This electrochemical energy is itself generated by proton transfer from the outside to the inside of the cell through membrane channels in the stator. These proton transfers cause the stator to distort its shape which generates a torque force on the rotor, causing it to rotate. With the rotor turning, movement is passed up the body to the hook resulting in a propelling motion in the filament.

This explanation is difficult to follow without a diagram of the flagellum, but lacking an image on the page just highlights the processes of trying to build a situation model in our heads when digesting the explanation. As a reader absorbs the description above, again referential inferences are being made, this time to concepts such as 'bacteria', 'cell', 'tail', 'filament', 'hook', 'motor', etc. More importantly, for one to understand the explanation causal inferences are made when constructing this model. These are not just reflecting what we are literally told about the model, but rather these inferences identify the properties we take various causal factors in the model to have in order that they do their causal work. For example, although it is not explicitly stated in the explanation, we have to make an inference about exactly how the distortion of the stator component is causally affecting a torque effect on the rotor. This could be by direct contact action, or by electrical forces operating across the gap between the two. To make sense of the model we have to fill this gap for ourselves. Another inference we have to make is about how proton transfer works. We are not explicitly told how the stator proteins permit the passage of protons from exterior to the cell wall, but some sort of process must be inferred if the model is to make sense.

I won't belabour the point with further examples, but from these few it should be clear that accommodating the causal and logical inferential steps we make while constructing mental (situation) models is essential to an adequate theory of scientific understanding. These cognitive steps take the agent beyond mere knowledge of an explanation, yet not so far as to entail the ability to solve prediction tasks. The inferential account I propose, although quite simple in its current form, accounts nicely for the difference between these levels of cognitive achievement. It also accounts for cases where it might be controversial as to whether we have understanding after all. The example of redshift hinted at this point, and the measurement problem and EPR paradox in quantum mechanics would further illustrate how the inferential account of understanding reflects how logical inference alone is often considered insufficient for genuine understanding.

There is much more to be done to develop this picture of scientific understanding. For a start, one very important area for further research is the metacognitive process of coherence evaluation of situation models. This is where we not only make inferences about components of our models, but we evaluate these inferences with respect to our background beliefs along dimensions of consistency and other coherence raising properties, such as simplicity and fecundity. The cognitive psychology literature is beginning to fill with studies that relate this sort of metacognitive processing with the less complex kinds of inferences we make directly in constructing mental models, and a philosophical theory of understanding must accommodate the results. This future research is as yet, however, beyond the scope of this article.

6. Conclusion

In closing I briefly wish to reiterate exactly what the inferential model of scientific understanding demands. According to this account, for an agent *A* to understand a phenomenon *p* described by an explanation using a scientific theory *T*, *A* must possess inferential knowledge of the reasons that are causally or logically responsible for the cause *C*, being the cause of *p* that it is, or for the logical properties *L* entailing the logical properties of *p* that they do. The agent therefore has to make inferences to the properties of *C* or *L* which are responsible for *p* being the way it is. These have to be the correct inferences, else *A* simply misunderstands *p*. They also have to actually have been made by the agent, else they constitute just potential understanding, not actual understanding. This is reflected by the agent actually constructing a situation model for *p* rather than merely knowing the explanation but not understanding it. This account, I believe, adequately captures both what we mean by saying someone understands a scientific explanation for *p*, and also the difference between knowing that explanation, understanding it, and using it to solve problems.

Acknowledgements

Thanks to two anonymous referees of this journal and Meghan Newman for helpful comments on this paper. Also thanks to Andrew Terjesen, Jason Ford, Sean Walsh, and Tristram McPherson for useful discussions on this topic.

Notes

[1] In this article I am specifically addressing scientific understanding, rather than understanding more generally construed. That there is a difference between the two will be argued shortly, and plays a crucial role in my argument.

[2] Trout is not offering these conditions as part of a semantic analysis, so these should not be taken as individually necessary and jointly sufficient conditions on understanding. They are a hedged description of the properties contingently accompanying understanding. Nevertheless, I argue they are an inaccurate characterization.

[3] On the other hand, one might argue that since theories in science often appeal to modeling techniques that deliberately incorporate idealizations and abstractions, we are asking too much of our theory of understanding if we require explanations be even approximately true. Elgin (2004) has argued that acquiring scientific understanding is often orthogonal, or even directly conflicts with acquiring truth. This would be correct if we didn't replace the aim of truth for that of approximate truth, but once this replacement is made, I believe her concerns are diminished.

[4] As one reviewer pointed out, most leading externalist philosophers offer a synthesis by supplementing their accounts of knowledge with an internalist 'no-defeater' clause. However, this fact does not show how an internalist condition like (2) can be reconciled with (3), since even with such a clause, one would still have to externalize the necessary 'supporting beliefs' in Trout's account.

[5] The idea of course is that 'mere' knowledge is the concept traditionally analysed in epistemology—something like justified true belief, and that understanding, we generally think, goes beyond this concept. Grimm (2006) provides a nice analysis of why understanding might still be 'mere' knowledge, arguing amongst other things, that Kvanvig's (2003) view must be in error. Kvanvig thinks understanding requires making the appropriate connections between our beliefs, whereas Grimm sees the focus more as being on our ability to answer questions. In this article I focus on the views of Trout and de Regt because they specifically concern scientific rather than general understanding.

[6] Bernoulli's principle states that for a fluid in an ideal state, pressure and density are inversely related, which entails that a moving fluid exerts more pressure when its velocity is low, and less pressure as the velocity increases.

[7] Notice in this explanation there is no mention of Bernoulli. That is because this explanation has been given in terms of pressure distribution with no reference to velocity distribution. Here's how we can relate the two: Newton's laws of mechanics can be expressed either as a

set of simultaneous differential equations or as a set of integrals. Using differentials for aerodynamics is useful if one wishes to determine the behavior of the fluid at specific locations or many locations while mapping the flow-field. This is where one would instead use Bernoulli's law, which is a version of the conservation of energy: $P_T = \text{constant} = P_S + \frac{1}{2}\rho V^2 = P_S + q$. This says that total pressure (P_T) equals static pressure (P_S) plus dynamic pressure (q), where this dynamic pressure is one-half fluid density (ρ) times velocity (V) squared. This latter approach to describing lift gives one the pressure distribution for a known velocity distribution. In order to determine this velocity distribution for the streamlines of the fluid one does however need to solve equations for conservation of mass, momentum and energy as the fluid passes the aircraft (Navier-Stokes equations). In simpler accounts at the macroscopic scale integrals that more closely resemble Newton's laws are used to describe changes in momentum and energy in regions of fluid flow. Both approaches are equally valid for describing lift, it is just that Newton's is simpler, and loses nothing provided one does not need to evaluate details of fluid flow. Also notice that use of Bernoulli's equation should avoid a popular but incorrect explanation of lift, which suggests conservation of energy entails that fluid travelling over the top surface of an airfoil must move faster than that moving below in order to 'catch up' with the lower air.

[8] I am not here committing to a strong distinction between scientific and non-scientific explanations. In fact I consider this a false distinction. All I require though is that even if explanations lie on a continuum between common sense and informed science, most explanations we find in science are far more complex than their common-sense counterparts.

[9] For an introduction to the discourse and comprehension literature, Kintsch (1998) and Tapiero (2007) are good places to start. Otero, Leon, and Graesser (2002) is a useful anthology regarding comprehension of specifically science texts.

[10] Notice that although in this literature comprehension is characterized in terms of knowledge, this does not lend much philosophical support to the claim that understanding is a form of knowledge. It doesn't undermine the claim either though.

[11] There are other means of representing the propositional content of an explanation of course, but here I illustrate merely that used most prominently in the literature.

[12] The last level of representation studied in the literature is that of *pragmatic communication*. This level reflects the primary message being conveyed by the explanation being given in the text. The pragmatic component of scientific explanation is well appreciated in the work of van Fraassen (1980) and Achinstein (1983), and can plausibly be thought to capture an important part of what it means to understand the world scientifically. It is less uncontroversial to suggest that this level of representation captures what we mean when we say we understand *p*. Consequently, I shall side-step the interesting question of how the pragmatics of explanation contributes to a cogent theory of scientific understanding. This covers the *levels* of representation, but representations themselves also come in many different *types*. I won't spend time explicating the wide range of representations used in the sciences, but at least a few recurrently play important roles: *class inclusion*; *temporal and spatial relations*; *composition of parts into subparts*; *step-wise procedures*; *causal chains and networks*; and *intentional action*. There is an important property of this list, and it is that the more fine-grained are these types of representation, and the greater their coherence (their conceptual interconnections), then the deeper is the knowledge one acquires of an explanation—or at least this seems to be indicated by studies (Graesser, Gordon, and Brainerd 1992).

[13] There are many ways of characterizing this taxonomy. For recent accounts, see Anderson and Krathwohl (2001) and Marzano and Kendall (2007).

[14] It is of course a common complaint amongst physics professors that students frequently fail to check that their answers cohere with common sense.

[15] The ordering I have given actually is traditionally thought of as being a ranking which originated with Bloom's (1956) taxonomy of educational objectives. I am not committed to it however.

[16] Like Kvanvig, I appreciate the importance of developing a coherent network of beliefs in order to achieve understanding. Unlike Kvanvig, I also think these beliefs must constitute knowledge.

[17] I omit the derivation because it requires graphics, which would consume too much space here. For a clear exposition of how to derive the angular radius of a rainbow, see Tipler (1991), 993.

References

Achinstein, P. 1983. *The nature of explanation*. New York: Oxford University Press.

Anderson, L. W., and D. R. Krathwohl. 2001. *A taxonomy for learning, teaching, and assessing*. 2nd ed. New York: Longman.

Bloom, B. S. 1956. *Taxonomy of educational objectives.* Handbook I*: Cognitive domain*. New York: McKay.

Cain, K., J. V. Oakhill, M. A. Barnes, and P. E. Bryant. 2001. Comprehension skill, inference-making ability, and their relation to knowledge. *Memory and Cognition* 29: 850–859.

Cote, N., S. R. Goldman, and E. U. Saul. 1998. Students making sense of informational text: Relations between processing and representation. *Discourse Processes* 25: 1–53.

Cushing, J. T. 1991. Quantum theory and explanatory discourse: Endgame for understanding? *Philosophy of Science* 58: 337–358.

Elgin, C. 2004. True enough. *Philosophical Issues* 14: 113–131.

Ericson, K. A., and W. Kintch. 1995. Long-term working memory. *Psychological Review* 102: 211–245.

Friedman, M. 1974. Explanation and scientific understanding. *Journal of Philosophy* 71: 5–19.

Graesser, A. C., and E. L. Bertus. 1998. The construction of causal inferences while reading expository texts on science and technology. *Scientific Studies of Reading* 2: 247–269.

Graesser, A. C., S. C. Gordon, and L. E. Brainerd. 1992. QUEST: A model of question answering. *Computers and Mathematics with Applications* 23: 733–745.

Graesser, A. C., M. Singer, and T. Trabasso. 1994. Constructing inferences during narrative text comprehension. *Psychological Review* 101: 371–395.

Grimm, S. 2006. Is understanding a species of knowledge? *British Journal for the Philosophy of Science* 57: 515–535.

Hempel, C. G. 1965. *Aspects of scientific explanation and other essays in the philosophy of science*. New York: Free Press.

Kintsch, W. 1998. *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.

Kvanvig, J. L. 2003. *The value of knowledge and the pursuit of understanding*. Cambridge: Cambridge University Press.

Marzano, R. J., and J. S. Kendall. 2007. *The new taxonomy of educational objectives*. London: Corwin Press.

McNamara, D. S., E. Kintsch, N. B. Songer, and W. Kintsch. 1996. Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction* 14: 1–43.

NASA (National Aeronautics and Space Administration). 2010. What is lift? [cited 28 August 2011]. Available from http://www.grc.nasa.gov/WWW/K-12/airplane/lift1.html; INTERNET.

Newell, A. 1990. *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Newell, A., and H. A. Simon. 1972. *Human problem solving*. Engelwood Cliffs, NJ: Prentice-Hall.

Otero, J., J. A. Leon, and A. C. Graesser. 2002. *The psychology of science text comprehension*. Mahwah, NJ: Lawrence Erlbaum Associates.

Regt, H. W. de. 2004. Discussion note: Making sense of understanding. *Philosophy of Science* 71: 98–109.

———. 2009. Understanding and scientific explanation. In *Scientific understanding*, edited by H. W. de Regt, S. Leonelli, and K. Eiger, 21–42. Pittsburgh, PA: University of Pittsburgh Press.

Regt, H. W. de, and D. Dieks. 2005. A contextual approach to scientific understanding. *Synthese* 144: 137–170.

Salmon, W. 1998. *Causality and explanation*. New York: Oxford University Press.

Tapiero, I. 2007. *Situation models and levels of coherence: Toward a definition of comprehension*. Mahwah, NJ: Lawrence Erlbaum Associates.

Tipler, P. 1991. *Physics for scientists and engineers*. Vol. 2. New York: Worth Publishers.

Trout, J. D. 2002. Scientific explanation and the sense of understanding. *Philosophy of Science* 69: 213–233.

———. 2005. Paying the price for a theory of explanation. *Philosophy of Science* 72: 198–208.

van Fraassen, B. 1980. *The scientific image*. New York: Oxford University Press.

Woodward, J. 2003. *Making things happen: A theory of causal explanation*. New York: Oxford University Press.