THIS IS A PREPUBLICATION DRAFT OF THE FOLLOWING PAPER:

KEIL, F. C. (2011). THE PROBLEM OF PARTIAL UNDERSTANDING. CURRENT TRENDS IN LSP RESEARCH: AIMS AND METHODS SERIES: LINGUISTIC INSIGHTS - 144, 251-276.

FRANK C. KEIL

The Problem of Partial Understanding

Abstract

The ways that people naturally build mental representations of explanations have important implications for how technical descriptions are crafted for broader audiences. All lay people, as well as experts outside of their specific areas of expertise, are limited to only partial understandings of complex causal patterns that underlie natural phenomena and many devices. These understandings are particularly sparse at the level of representing specific mechanisms creating a need for supplementing their sparse understandings with ways of accessing experts when needed. A central issue concerns illusions in which people tend to overestimate their levels of mechanistic understanding such that they often may not fully realize the gaps in their understanding and when to ask further questions of experts. Although such illusions can cause some complacency with respect to one's own explanatory understanding, people of all ages have other skills that enable them to track more abstract causal patterns they do given them sense of fertile areas of expertise and how to link specific phenomena with particular groups of experts. A better

understanding of the cognitive science of these skills is an important resource to those concerned with languages for specific purposes.

Introduction

Explanatory understanding, namely one's understanding of how things work or why a phenomenon occurs as it does, is important to all individuals from early childhood on (Keil, 2006). While individuals may differ in the depth of understanding that they seek (Cacioppo, Petty, Feinstein, & Jarvis, 1996), everyone seeks to grasp at least a partial reason for a pattern, usually in causal terms. But, even though everyone seeks causal understanding to some degree virtually no one seeks it exhaustively in any instance. If people normally have such gaps in their understanding, it would be useful to those concerned with technical languages to know more about those gaps and how people cope with them, How much of the real world causal structure and relations can an individual layperson grasp and how well does one know one's own gaps? A closely related question concerns whether one's coarse understandings represent a reasonable amount to grasp given the nature of one's cognitive limitations and given the nature of alternate means and sources for further grounding one's understandings. Thus, when crafting explanations for audiences beyond the most knowledgeable experts, it is helpful to get a clearer sense of what sorts of coarse causal representations are the most cognitively natural and how people supplement gaps in their own minds by accessing other sources of knowledge. Any time one attempts to paraphrase a technical description of a complex phenomenon for a broader audience one has to intuitively prune out some details on the assumption that those details are not central to gaining a lay understanding. It is useful therefore to have some sense of what lay understandings typically look like, that is what it means to have a coarse yet functionally effective understanding. Here, I suggest that a full characterization of such coarse understandings must include a description of how laypeople learn to “outsource” some complex details by deferring to more expert groups and how they know their own knowledge shortcomings.

Learning to defer and supplement our own coarse understandings, however, requires some sense of the limits of our own knowledge . In fact most people do not have an accurate sense of how well can they explain a phenomenon. They do not do a good job tracking how well they have mentally represented the workings of a system and of the extent to which there are still gaps and limits to their understanding. This creates a challenge for those interested in creating explanations for others, as the recipients of such explanations may often overestimate how much of an explanation they been able to represent mentally. It is therefore important to have new methods for assessing explanatory understanding and of meaasuring the mismatch between what people think they know and what they really know. For that reason, one such method is a major focus of this paper.

If people attempt to handle gaps in understanding by deferring to others who can then provide information to fill in those gaps, it is important to better understand the cognitive underpinnings of deference. Deference is commonplace in all cultures as cultures themselves are partially defined by the divisions of labor, and of corresponding cognitive labor, that help them give added value to each of their members (Durkheim, 1893/1997; Smith, 1776/1937). Do people, however, defer when they should and in ways that are effective? It is quite possible that people sometimes don’t always defer when they should and that other times they defer when we should not. They might mistakenly defer to the wrong group of experts because they rely on the wrong sorts of cues to determine appropriate areas of expertise. A related question asks what sorts of factors govern gullibility. Under what conditions, and why, are people especially gullible? In addition are certain kinds of individuals, such as children, especially gullible in certain ways? Indeed, it has been suggested by some that children are in essence almost infinitely gullible (Dawkins, 1993). Although gullibility can be abused by those attempting to engage in propaganda or unethical marketing it can also be a valuable phenomenon to understand for those concerned with trying to convey information in ways that allow lay people to make informed judgments about that information.

In this article I will suggest that recent research is beginning to unpack the answers to these questions and that some of the answers may be counterintuitive in terms of where people do and do not track causal patterns and in terms of their awareness of what patterns they are grasping directly and indirectly through such means as deference. All people have severe limitations in terms of the amount of and kinds of information they can understand, in terms of insight into their own limitations, and in terms of their gullibility and deference. Given all those limitations, the question naturally arises as to how people ever cope with their explanatory limitations and whether their can ever effectively tap into expertise that is in other minds.

I have indicated that people have quite coarse understandings with substantial gaps. To better understand these gaps, it is important to characterize what natural explanatory understanding normally looks like at various levels of description. Consider, explanatory understanding construed as having mental blue prints of the world, how-things-works diagrams in the head. It is often held that novices have more concrete understandings and one version of concrete representations is something like mental models of a system's components and how they work together Such models might have fuzziness in terms of their precise details (e.g. not visualizing the wiring harness of a car) but still might be simplified concrete cartoonish mechanistic models; models that are like the neat, elegant oversimplified diagrams of how things work that are presented to children in school. Such models would sacrifice some detail for coherence, concreteness, visualizability, and clean predictability. This is one sense of idealizations in science – sometimes wrong and giving the wrong predictions, but overall useful because of their compact elegance (Bokulich, 2009). Consider, for example, the central dogma of molecular biology, or models of the solar system as a series of circular orbits. Do laypeople tend to do the same thing? Do they strive for mechanistic understandings with some simplifying idealizations that make them easier to visualize? Here, I will argue that they do not. Instead, they build understandings in ways that are quite different from our traditional understandings of mechanisms, but ways that are nonetheless effective. A different level of description of explanatory understanding is needed to understand this kind of competence without mechanistic details.

This article considers these issues by first considering in some details illusions of explanatory depth, that is the ways that people do indeed think they understand things in more detail than they really do. It then considers related illusions of insight, in which people think they have gained explanatory insight when they really have not, illusions that can seriously impair their assessments of whether they are building a good understanding as they listen to an explanation. These two illusions also reveal the ways in which our mechanistic understandings are limited. Such limitations in turn lead to alternative ways that people do make sense of the world without such details. I will suggest that they do so by tracking causal patterns at other levels that are quite distinct from clockworks mechanisms but that nonetheless can be effective guides towards "knowledge outsourcing," that is by tracking highly abstract causal patterns far above the level of mechanism that can nonetheless serve as cues to appropriate domains of experts for the phenomena at hand. I will further suggest that the abilities to both detect and use these patterns are foundational and early emerging in development and have strong implications for how laypeople understand divisions of cognitive labor in all cultures, and how laypeople cope with technical information. Thus one way of coping with technical information is not trying to master it fully oneself, but knowing which kind of expert would be most useful for unpacking particular details of interest and rendering them into appropriately simple formats.

Illusions of Explanatory Depth

I have suggested that understanding can take several distinct forms even as there is a tendency to equate it with having schematic quasi-visual diagrams in the head. In addition to such concrete representations, one can also have functional understandings of a system without any mechanistic details, such as knowing the functional components of a computer without having any idea of how those functions are mechanistically implemented. One can also have an even more abstract sense of causal powers of parts of system without clearly grasping functions. For example, one might know that magnets attract iron without any sense of either mechanism or of function. (For most non-living natural phenomenon, functional interpretations do not make sense). Finally, one might simply know that some kinds of properties are likely to be causally central to some domains without knowing any further details, such as knowing that surface colors tend to be more intimately causally connected to other properties of natural kinds than they are to artifacts. (We assume that colors of flowers are more important to understanding what the flower is than colors of chairs to chairs). Despite having understandings at all these levels, people tend to focus on the mechanistic level and it is here that there may be the greatest mismatch between what we know and what we think we know.

There is now considerable evidence that people often think they understand complex phenomena more "deeply", that is with greater precision and depth, than they actually do. This effect is found in both adults (Rozenblit & Keil, 2002) and children (Mills & Keil, 2004) and across many content domains ranging from various artifacts, to natural phenomena (Rozenblit & Keil, 2002) to social, political and economic phenomena (Alter, Oppenheimer, & Zemla, 2010). We have called this error "The Illusion of Explanatory Depth", or IOED.

The illusion is typically documented by teaching a group of participants how to use a scale corresponding to their level of understanding of a device or phenomena. For example, we presented them with a series of diagrams and accompanying text indicating what a level 7 to level 1 understanding of a device such as a crossbow might be. A level 7 understanding is essentially knowing every part of a device, all its properties, how all the parts fit together and their properties as a relational whole in terms of overall functions and functions of all coherent subunits. Put differently, if one had all the raw materials and an appropriate machinist or craftsperson at one's side, one could essentially construct the device from scratch. A level 4 understanding is one where one has some details but only in a schematized outline sense that might be missing many details, including quite possibly some critical ones. A level 1 understanding would correspond to having a vague impression of what the device looked like in silhouette-like manner and having perhaps a crude sense of either its causal power or overall function. For example, a level 7 understanding of a crossbow would entail knowing every single component and why the components were the ways they were, including each component's shape and material composition, knowing the functions of all the subunits such as the trigger mechanism, and the crank mechanism, and then knowing how they all fit together to give rise to the function of launching arrow-bolts with great force and speed. A level 4 understanding might involve knowing that there needs to be a trigger and some way of building up and storing a huge string tension and that such mechanisms enable launching arrow bolts at large speeds. A level 1 understanding might simply know the overall shape of crossbows and know that they are used for shooting, or perhaps even more weakly that they are some kind of weapon. Participants in these studies were presented with each level of understanding using the crossbow as an example and were told that they would be asked to rate how close their own understanding was to one of these seven levels. Thus, before they engaged in any ratings of their own knowledge they were fully familiar with the idea that they were to use the scale to indicate how well they thought they understood each device or phenomenon.

Having been carefully trained on such a scale, participants are then asked to rank, relatively quickly, their level of understanding of a large list of everyday items (e.g., a 48 item list). We'll call these initial ratings

R1. Once they are done with such ratings, they are then told to focus on just four items from that list. Different participants are asked to focus on different subsets of such items to ensure that effects are not being just caused by a few specific items. Participants are then asked to write as full explanations of each for the four items as they can, and then, in light of those explanations, to re-rate their initial understanding (R2). They are then given asked to answer a critical test question for each of the four items. The test questions are designed to get at the heart of a deep understanding of the object in question (e.g. how does a helicopter go from hovering to flying forward?). Then, in light of that answer, they are asked to again re-rate their initial understanding (R3). Next, they are given a concise but detailed and accurate expert explanation of each of the four items and then, in light of that explanation, to re-rate their initial understanding (R4). Finally, to show that any potential drops in ratings are not merely caused by a general decline in confidence, participants were asked to rate their final level of understanding after having studied the expert explanations (R5).

Studies of this sort consistently show a large drop in ratings from the initial rating (R1) to R2, a further drop to R3, essentially the same low rating at R4 and then a surge to the highest rating R5, after studying an expert explanation. Thus, the findings show both a strong mismatch between what people think they knew and what they later acknowledge that they really knew. They often displayed marked surprise at their own levels of ignorance having been quite convinced that they had a much clearer idea of how something worked than they really did. In addition, because the ratings were uniformly high at the end, it is possible to discount the idea that participants were simply having their confidence shattered and became unable to give any high ratings of their understanding in any condition. Whether the explanation is about how a device such as a helicopter works, about how a biological entity such as a kidney functions, or even how a natural phenomenon such as a thunderstorm forms, an IOED is found.

The Specificity of the Illusion

The IOED is quite specific to explanatory understanding. When people rate their quality of their knowledge in other areas, they tend to be quite a bit better calibrated either not showing any illusions at all, or much more modest ones. Consider the cases with facts, procedures and narratives (Keil & Rozenblit, 2002). Each of these cases showed a clear contrast with explanatory understanding. In particular, although people strongly overestimate the depth and detail of their explanatory understandings they seem to be much better able to inspect the quality of their knowledge in these other areas, even in cases where that knowledge is at the same level of accuracy as explanatory knowledge. It is informative to examine each case in more detail.

Facts

When people are asked to rate whether they know a fact, they tend to be well calibrated. For example, when participants were asked to rate how well they knew the capitals of various countries, their confidence in how well they knew each capital from a long list of countries reflected their performance and their later revised judgments of their initial knowledge. Thus, their R1 ratings and their R2 ratings were very similar and didn't show a large drop. One likely reason for this accuracy is that it is much easier and quicker to self-test one's knowledge of facts than it is to test one's understanding of explanations and one can more easily tell if one actually has successfully retrieved a fact as opposed to a full explanation. The accuracy for self-estimates of factual knowledge was found even though the overall accuracy was only moderate, as many of the capitals were obscure.

Procedures

People's accuracy with facts may occur simply because facts are usually brief and therefore easy to check in real time. Knowledge of procedures is more complex and may involve quite complex series of steps. Even so, procedures differ from explanations in that one might have clearer knowledge of whether one has engaged in that procedure successfully in the past. We provide full explanations much less often, even for those phenomena where we may actually possess a full understanding. To explore this possibility, participants in another study were asked to quickly rate how well they knew how to do various procedures, such as making an international phone call, or baking a layer cake. Again, as for facts, participants were much better calibrated than for explanations, often showing little or no drops in ratings between R1 and R2, where they had to write out a rating. Thus, the illusion for explanation is not simply because the information is more extended than for facts. It is an open question as to whether the more tacit nature of some forms of procedural knowledge makes it somehow easier to self-assess as that dimension has not yet been specifically studied. It is clear, however, that at least some forms of procedural knowledge tested were quite explicit in nature, such as knowing the steps needed to place an international phone call. For many citizens of the United States, at least, they make such calls quite rarely and therefore have not encoded such actions as an automatic implicit procedure.

Narratives vs. Explanations

Narratives would seem to be a good bit closer to explanations in that they don't usually have to be performed in order to be learned, in contrast to procedures. Moreover, people sometimes describe theories as kinds of stories. However, it is much more common to tell narratives than explanations. In addition, explanations may have several features that provide some sense of insight that are not present in narratives. For example, explanations can exist at both higher functional levels and at lower more mechanistic levels. One can "understand" how a computer works by knowing how to turn it on, use the keyboard and mouse and adjust various controls. This functional knowledge might be accompanied with a sense of mastery that might seem like explanatory understanding, a sense that might be partially confused with mechanistic understanding. For narratives, such function-for-mechanism confusions do not normally occur. When adult participants are asked to rate how well they know the plots of various movies, their initial ratings (R1) do not differ that much from their ratings (R2) after they attempt to write out the plots. Apparently, even though narratives have surface similarities to explanations in terms of their coherence and some similarities of discourse structure, they also differ in important ways that prevent them from causing strong illusions of knowing. Thus, function-for-mechanism confusions may simply not occur with narratives. Second, because people tend to have much more practice providing narratives of familiar life events they simply might have less experience with encountering gaps in their explanatory knowledge and may therefore remain miscalibrated.

The specificity of the illusion of explanatory depth holds for children as well as for adults (Mills & Keil, 2004). In addition, preliminary studies suggest that it holds for ratings of peers as much as it holds for ratings of the self. This lack of a self-other difference suggests that the illusion does not arise from effects known as self-enhancement biases (Krueger & Dunning, 1999). A strong majority of people tend to think they are above average on most skills, even though that is a mathematical impossibility. If those self-enhancing biases were the only reason behind the illusion of explanatory depth, people should not show the illusion when rating peers. The persistence of the illusion when evaluating others suggests that factors such as the function-for-mechanism confusion are at work as the confusion would hold for both the self and others. Interestingly, if one asks people how well they could learn about a phenomena, not how much they actually know, the self-other differences start to emerge, with more learning ability attributed to the self. This finding makes sense since such intuitions about learning ability do not involve inspection of one's current knowledge and making misattributions about such things as functional vs. mechanistic understanding. Instead, these intuitions rely more on person's sense of their general intelligence, which is well known to have strong self-enhancement biases (ibid). In terms of individuals confronted with technical language, this means that people might overestimate their ability to master a difficult explanation before really studying it, which is in contrast to the IOED in which people area estimating their existing levels of understanding.

Reasons for the IOED

I have already alluded to some reasons why the IOED may be much stronger for explanatory understanding. There is much more potential for function-for-mechanism confusions with explanations such that when one does gain a legitimate insight into a functional relationship, one might confuse that insight with a deeper mechanistic understanding. The IOED may also occur because we rarely attempt to give full explanations and therefore have much less experience with our abilities to provide them to others. An additional factor may be because we become confused by our ability to construct detailed mechanistic explanations "on-the-fly" when a device or other object is in front of us. We may mistake our abilities in such cases with having a full representation of the mechanism in our heads. For example, if one were given a stapler and asked to explain precisely how it works, one might do so perfectly, and assume from that success that one had all that knowledge in mind before ever being handed the stapler. Yet, if asked to do so without the stapler, one might have serious gaps in understanding. In a similar way people often think they have remembered great details from visual scenes only to turn out having very incomplete memories (Levin, Momen, Drivdahl, & Simons, 2000). Because people can so easily revisit most scenes with their eyes, they often overestimate how much they actually are storing of the scene when they are not looking at it. These factors, as well as others, may all converge to create an especially powerful illusion of knowledge depth for explanations. Another factor that may contribute to the illusion is our surprising ability to keep inconsistent or even contradictory ideas in mind about the same topic without recognizing the contradiction (Epstein, Glenberg, & Bradley, 1984; Glenberg, Wilkinson, & Epstein, 1982). Thus, when gaps in understanding might point out an inconsistency or flaw, we may simply not notice it.

Illusions of Insight

The illusion of explanatory depth may be exacerbated by a related illusion that we will call the Illusion of Insight. This illusion occurs when we are hearing an explanation and an extraneous irrelevant bit of information nonetheless causes an apparent, but false, rush of insight. If explanations are more susceptible to such false revelations, then they would be more susceptible to illusions of explanatory depth. One factor that may contribute to such illusions of insight is the use of highly concrete information in the context of trying to understand less tangible patterns. Consider, for example, the use of brain imaging to shed insight onto psychological phenomena. There has been an explosion of research in cognitive neuroscience, much of it reflecting real advances in using brain structures and processes to constrain psychological models. Yet, there may also be cases where the simple use of brain imaging gives one a sense of having a better grasp of a phenomenon when one really does not.

To test whether brain-imaging findings could cause illusory explanatory insights, we designed a study in which participants heard one of four types of explanations for several phenomena and were asked to rate the quality of the explanations. The four types were: good psychological explanation, bad psychological explanation, good psychological explanation plus irrelevant imaging information and bad psychological explanation plus irrelevant imaging information. (Weisberg, Keil, Goodstein, Rawson, & Gray; 2008). For example, one phenomenon was the “curse of knowledge” in which people tend to assume that others know the same sorts of details that they do about a situation even when they have different background experiences. A good psychological explanation described a mechanism in terms of difficulty in changing points of view and thereby mistakenly projecting one's own perspective on others. A bad psychological explanation merely paraphrased the phenomenon in different terms without offering any new mechanistic details. The irrelevant neuroimaging information merely reported that "brain scans" indicated that the curse happened because of " "frontal lobe brain circuitry" already known to be involved in self-knowledge. No images or graphs of neuroimaging were necessary. The bad explanations were circular in nature and simply restated the phenomena without really shedding insight. Participants found it easy to tell good explanations from bad ones when they did not contain neuroimaging results. The neuroimaging data was also irrelevant and easily judged to be so by neuroscience experts. But, for those who were less experienced in neuroscience, the mere mention of neuroimaging findings, even when it was clear that they were not adding any new information, was enough to sharply raise the ratings of the quality of the bad psychological explanations, often to a point where they were not distinguished from the good ones.

This finding and related ones with actual images (McCabe & Castel, 2008) demonstrate that a sense of acquiring an explanatory insight cannot always be justified and that certain kinds of information can "seduce" one into thinking one has gone deeper in explanatory understanding when one hasn't. Neuroimaging data may be especially seductive because its common way of being depicted is often confused as literal photographs of the brain's activities when in fact it is really a result of complex computations and inferential statistics (Roskies, 2007; 2008). It is easy to see how experiences with seductive data of this sort might enhance the IOED.

It seems likely that illusions of insight will be found in many other areas where a form of data offers the tantalizing hope of making more concrete a less tangible form of explanation. Attempts to claim that a particular psychological phenomenon can be explained by a "gene" for that phenomenon may also provide false rushes of insight when the genetic evidence is in fact not really informative. Similarly, when amorphous and ill-defined categories such as cancer are explained by discrete single causes, a false rush of insight may occur even when that single cause is non-informative. There is a need for systematic studies that try to develop a more principled characterization of under what circumstances illusions of understanding contaminate our ability to sense how much of an explanation we have really grasped.

Interim Summary

People frequently labor under delusions of explanatory competence when in fact they may understand very little about a phenomenon or device. In addition, they may often build up a sense of explanatory competence by false illusions of insight. In many cases their real knowledge can be shockingly limited. For example, in one study roughly half of adult British subjects were unable to recognize severe flaws in a bicycle drawing that would make the bicycle completely unable to be actually ridden (Lawson, 2006). It might then seem that, outside of very narrow areas of expertise, people's explanatory understandings are severely limited. In fact, people do track causal patterns in the world in surprisingly effective and subtle ways, just not at the level of detailed mechanisms. In what follows I briefly indicate what they do track in terms of causal patterns and how that information can form an effective means for buttressing their extremely sparse knowledge of mechanism.

Using Coarse Knowledge Effectively

What do people know if they do not know mechanisms in much detail? They develop an extraordinary sense of patterns in high level domains such as folk biology (Atran & Medin, 2008), folk physics (Bertamini, Spooner & Hecht, 2004) and folk psychology (Sodian & Kristen, 2010). It is beyond the scope of this paper to describe what they know in much detail (see Keil, 2010); but there is strong evidence that they track many patterns, such as what domains are likely to have a great deal of causal structure as opposed to minimal causal complexity, what domains have entities that have overall functions for their objects (e.g. tools can be "for x") and what ones do not (e.g., animals cannot be "for x" even if their parts, such as a shell, can be for something such as protection). They know that in some domains, such as artifacts, shape usually matters a great deal for understanding their nature whereas in other domains (such as plants) color might also be important (Keil et al., 1998: Keil, 2010). They can know that certain entities can have particular classes of causal powers even if they do not know how those powers work. For example, they can know that intentional agents are the usual causal mechanisms behind events that change an array from a disordered one to an ordered one (Newman et al., 2010).

All of these would be impressive things to know as adults. The remarkable finding is that they are also known by young children, and sometimes even by preverbal infants. For example, 12 month old infants clearly expect that only intentional agents can cause a system to go from a state of order to disorder (ibid). Similarly, well before the start of schooling, children think that it is more reasonable to ask what a tool as a whole is for than what an animal as a whole is for (Greif et al., 2006). Thus, throughout much of a child's development, there is effective tracking of a great deal of information about high level causal patterns associated with broad domains, even as all sorts of details are missing.

Children and adults alike use this information about causal patterns in several ways to further their understanding on the world. One such way is to be able to construct explanations "on the fly". While they may not have a fully detailed understanding in advance, when they are immersed in a situation, that coarse causal knowledge and sense of high-level patterns seems to guide explanations in the moment. Thus, a person who figures out in real time how a bicycle derailleur works may effectively have a fully complete clockworks knowledge of the device that emerges through the interactions of general senses about mechanistic systems and functions combined with the specific details of the object in hand. Only a few moments later, however, when that person leaves the scene of the derailleur, the mechanistic model may quickly fade as the object-supplied support is no longer present. It is simply too taxing to try to keep all those details in mind for each and every device that we encounter and instead is far more efficient to have just enough information to be able to decipher them when we are confronted with them.

Many detailed mechanistic explanations therefore exist as cognitive ephemera that are critically supported by the presence of an object or phenomenon for inspection and manipulation and not as purely mental representations with all the details when the explanandum is not present. People may draw from their successes at constructing detailed explanations in such situations the erroneous inference that they have internally mentally represented all the details when they really haven't (hence the IOED), but they still have succeeded nonetheless in that situation. The problem is that they may infer from their success in a task with an object present that they have a fully formed mental representation that they could use in other contexts that transcend that situation. Thus, they might erroneously think they could explain a device to others without having it on hand. Or they might falsely assume that they could fully understand another person talking about some new detail of how a device worked even though they do not have a diagram of the object present to supplement the verbal explanation. Similarly, a person providing an explanation might erroneously assume that others have far more detailed internal representations than they really do and thereby not bother providing concrete information that is essential for constructing a full understanding.

Much of the time, however, people make progress in acquiring explanatory knowledge while not having much in the way of specific understandings. Consider how the causal patterns that people do easily track may be helpful in choosing between competing explanations even when one doesn't fully understand or remember the details. For example, when hearing two alleged experts explain the workings of an unfamiliar tool, adults and children alike will tend to prefer the explanation offered by an expert who focuses on shape, strength and size and not on color, surface patterns or precise numbers of parts (Keil, 2010). They will show such preferences even though they do not have information about the detailed workings of the tools. It is therefore possible to prefer what is likely to be good explanation for particular entity without much understanding of its details. One merely needs to know that certain kinds of properties are more likely to be causally central in some domains (e.g., tools) than in others (e.g., animals).

Similarly, high-level causal patterns may help guide inferences about how to construct appropriate gists of larger more complex explanations. A critical challenge for those working with technical language is to figure out ways to compress it into forms that are more accessible to more people and to do so in ways that capture the key ideas without major distortions. A sense of the high level causal patterns that are essential to a domain may help people focus on those parts of an explanation that refer most often to those causal patterns, just as causally central features heavily influence categorization (Sloman, Love, & Ahn, 1998). Although one can use strategies such as focusing on topic sentences and repeated elements to infer the most central elements for a summary of a complex explanation (Marcu, 2000); people can also use more subtle structural principles to derive the most appropriate gists, principles that often build off their abilities to track high level causal patterns in specific domains, such as looking at how certain elements elaborate on others in ways that form nested hierarchies of causal elaborations (Rottman & Keil, draft). The most important use of such patterns, however, may be to guide people to appropriate routes of deference to other minds.

Deference and the Division of Cognitive Labor

Traditionally much of the research on people's abilities to understand the world around them has been conducted as if each person is an isolated individual attempting to figure how things work without any assistance from others. This is, of course, far from the truth. In practice, the majority of our explanatory understandings are deeply dependent on knowledge in other minds (Gelman, 2009). To understand almost any topic in contemporary science requires many simultaneous acts of deference to different groups of experts. A person who is diagnosed with diabetes and wants to further understand the disease cannot hardly hope to understand all the mechanisms in full detail that are responsible for both normal physiology and for the disruptions associated with that disease. Instead, that person must try to construct a schematic understanding full of pointers to various subareas of expertise. We all do this so automatically and frequently that we can overlook the rich cognitions that underlie and support the acts of deference themselves.

One of the first accounts of the importance of deference was formulated by Putnam (1975) in his discussion of the meanings of natural kind terms. Putnam argued that meaning of terms is not simply in the head but also arises from deference to experts. Thus, one might oneself know the difference between platinum and silver but would believe that there are experts who could tell such a difference and that, by virtue of knowing that those experts exist, one indirectly "knew" the meaning as well. The same idea of deference expands far beyond the meanings to terms to almost all our explanatory understandings of complex phenomena and devices. We are constantly outsourcing parts of our understanding to more expert communities and feeling some confidence about own understanding by having reliable chains of deference that we can use.

A series of recent studies shows that not only adults, but also young children are capable of navigating the divisions of cognitive labor that are intrinsic to all cultures (Danovitch & Keil, 2004; Keil, Stein, Webb, Billings, & Rozenblit, 2008; Lutz & Keil, 2002). One way to assess this ability is to present children with problems of the following form:

> You want to know why people sometimes fight more when they are tired. Who should you ask to help you understand this?
>
> A.Someone who knows all about why people smile at their friends when they see them.
>
> Or
>
> B. Someone who knows all about why salt on people's icy driveways makes the ice melt sooner.

Even kindergarten age children are often able to judge at above chance levels that a person such as expert A is more likely to be helpful. They do so even though they clearly know very little about psychology or chemistry. Instead, young children figure out the appropriate areas of expertise by using their coarse tracking of causal patterns (Keil et al., 2008). Thus, they might know that patterns in one domain are primarily those arising from action at a distance (social phenomena) as opposed to direct contact (physical mechanics) and use that contrast to narrow down the set of appropriate experts. Using those sorts of causal patterns, in conjunction with various ways of evaluating the veracity of sources (Mills & Keil, 2005), enables them to often lock onto the appropriate groups of experts even when they may directly know very few explanatory details themselves. These skills are never perfect, especially in younger children who can be misled into deferring to the wrong expert, but they are strikingly accurate in many cases and help them leverage their own modest understandings.

Children are also able to look at various phenomena and decide if they are likely to be areas of fertile expertise in general. For example, they can be quite good at sensing if a category of things is likely to be one that requires expertise to understand or not, that is whether it has enough causal density to warrant an expert community. Consider, for example, if a child is asked whether it is more plausible to have an expert on "Dogs with red collars" or "Dogs that hunt." By seven years of age most children will judge that it only makes sense to have experts on dogs that hunt as the other category is simply structured by one criterion and has no internal structure. Insights like this can greatly help young children, and people of all ages, to know what sorts of phenomena to focus on further so as to figure out appropriate supporting communities of experts.

Children and adults alike need to have some sense of legitimate areas of explanatory expertise, areas that track complex causal phenomena and provide insight into them. There are other kinds of experts, such as trivia buffs who may know huge quantities of facts about a domain without any explanatory underpinnings, such as a fan who knows the favorite colors of every member of a sports team, or the astrologer who knows vast details about alleged alignments of celestial bodies but is mistaken about any connection between such details and the courses of people's lives. People are not always able to discern the correct areas of expertise (witness the continued success of astrology); but they often can sense where there are rich causal patterns for which there is a legitimate need for expert explanations. There is a need to better understand the conditions under which people do succeed at discerning appropriate areas of expertise. Those who craft technical explanations and paraphrase them for broader audiences might then be able to better supply valuable clues to when an area of expertise is legitimate.

Conclusions

It is impossible for any one person to directly have complete explanatory understandings for most natural and artificial entities, even if those explanations are fully known in the appropriate expert communities. At best, one might have extraordinary depth of understanding in one, or possibly a few, narrow areas. This creates a challenge. How are people, at any age, able to make some sense of the world around them when they clearly cannot carry within their own minds all the details? It is important to understand this process so as to be better able to craft explanations that fit with the ways that people naturally do construct coarse representations and then develop networks of deference to support their gaps in their own understandings.

I have argued that adults have a well-developed set of strategies for tracking abstract patterns in domains that are represented far above the level of concrete mechanisms. These more abstract patterns and skeletal schemas serve two roles: 1. They enable people to grasp in a crude sense various entities well enough to react to them (if they are natural) or use them (if they are artifacts) and 2. They enable people to navigate the division of cognitive labor by focusing in on appropriate groups of experts to ground their incomplete and gap-laden understandings. These skills are so foundational that they are found in young children. Such skills are also open to considerable improvement through the lifespan. Education can enable people to become somewhat more detailed in their understandings and it can be a major challenge to learn how to pick up subtle cues as to legitimate areas of expertise when there are several competitors for one's attention. An important area of future research concerns ways in which one might teach people to further sharpen these skills.

One major challenge concerns the phenomenon that everyone labors under illusions of understanding, thinking that they can explain things much better than they really can on their own. Such illusions, however, may not be as detrimental as they seem. Although they do generate false feelings of knowing and insight, they may in fact be better proxies for knowing in the more indirect sense of having access to the appropriate experts. It may, in fact, be adaptive to not keep diving deeper in a domain to gain full understanding when a high level gloss will work much of the time and a chain of deference is available when more details are needed. The illusions of knowing may be a helpful way of keeping people from getting unnecessarily lost in a sea of details. The task for the future is to better understand how effectively people do use these pathways to expertise and how well experts cooperate in this enterprise.

All of these issues should be of central concern to those interested in communicating technical language to broader communities and to those wrestling with how to best paraphrase complex explanations. If most people seem to naturally settle in a certain range of coarse mechanistic understandings, it is important to understand the average level and the extent to which one can guide people to more detailed understandings in that range and when one goes too far and overwhelms almost anyone with excessive details. It is equally important to know that the goal of explanations is not merely to convey understanding at the level of mechanistic details. One might also confer major benefits by helping a person grasp higher order patterns of a phenomenon, such as that it involves a positive feedback loop or that it exhibits exponential growth., or that it is composed of many repeated units that work the same way in parallel. Cognitive scientists are only just starting to consider the ability to discern and remember these kinds of causal patterns (Kim et. al, 2009). A better understanding of such abilities would enable one to know how to more effectively direct people towards appropriate communities of experts through highlighting causal patterns that they can grasp and showing how such patterns are connected to experts. It is similarly important to know the ways in which people might be misled by their overestimation of their own levels of understanding. Knowledge of those illusions might enable one to construct an explanation with cues or checks that will enable a reader to better know when they have gaps. Taken together, there is enormous potential in future collaborations between these areas of cognitive science and scholars concerned with languages for specific purposes.

References

Alter, A. L., Oppenheimer, D. M., & Zemla, J. C. (2010). Missing the trees for the forest: A construal level account of the illusion of explanatory depth. *Journal of Personality and Social Psychology*.

Atran, S., & Medin, D. L. (2008). *The Native Mind and the Cultural Construction of Nature*. Boston, MA: MIT Press.

Bertamini, M., Spooner, A., & Hecht, H. (2004). The representation of naive knowledge about physics. In G. Malcolm (Ed.), *Multidisciplinary Approaches to Visual Representations and Interpretations*. Elsevier.

Bokulich, A. (2009). How Scientific Models can Explain. Synthese, 171; 1-13

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: the life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*, 197–253

Danovitch, J. and Keil, F. C. (2004). Should you ask a fisherman or a biologist?: Developmental Shifts in Ways of Clustering Knowledge. *Child Development, 75*, 918-931.

Dawkins R. (1993). Viruses of the mind. In Dahlbom B. (Ed.), *Dennett and his critics: Demystifying mind* (pp. 13–27). Oxford, England: Blackwell.

Durkheim, E. (1997). *The division of labor in society* (G. Simpson, Trans.). New York: The Free Press. (Original work published 1893)

Epstein, W., Glenberg, A. M., & Bradley, M. M. (1984). Coactivation and comprehension: Contribution of text variables to the illusion of knowing. *Memory & Cognition, 12*, 355–360.

Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology, 60*, 115–140.

Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory & Cognition, 10*, 597–602.

Greif, M., Kemler-Nelson, D., Keil, F.C. & Guiterrez, F. (2006). What do children want to know about animals and artifacts?: Domain-specific requests for information. *Psychological Science, 17*, 455-459.

Keil, F. C, Smith, C. S., Simons, D., & Levin, D. (1998) Two dogmas of conceptual empiricism. *Cognition, 65,* 103-135.

Keil, F. C. (2006). Explanation and Understanding. *Annual Review of Psychology, 57*, 227-254

Keil, F. C., Stein, C., Webb, L., Billings, V. D., & Rozenblit, L. (2008). Discerning the Division of Cognitive Labor: An Emerging Understanding of How Knowledge is Clustered in Other Minds. *Cognitive Science, 32*, 259-300.

Kim, N. S., Luhmann, C. C., Pierce, M. E., & Ryan, M. M. (2009). The conceptual centrality of causal cycles. Memory & Cognition, 37, 744-758.

Krueger, J. & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*, 1121–1134.

Lawson, R. L. (2006). The science of cycology: Failures to understand how everyday objects work. *Memory and Cognition, 34*, 1667–1675.

Levin, D. T., Momen, N., Drivdahl, S. B. & Simons, D. J. (2000) Change blindness blindness: The metacognitive error of overestimating change-detection ability. *Visual Cognition, 7*, 397–412.

Lutz, D. R., & Keil, F. C. (2002). Early Understanding of the Division of Cognitive Labor. *Child Development, 73*, 1073-1084.

Marcu, D. (2000). *The theory and practice of discourse parsing and summarization*. Cambridge, Mass.: MIT Press.

McCabe, D. P., & Castel., A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition, 107*, 343-352.

Mills, C., & Keil, F. C. (2004). Knowing the limits of one's understanding: The development of an awareness of an illusion of explanatory depth, *Journal of Experimental Child Psychology, 87*, 1-32.

Mills, C., & Keil, F .C. (2005). The Development of Cynicism, *Psychological Science, 16*, 385-390.

Newman, G. E., Keil, F. C., Kuhlmeier, V., & Wynn, K. (2010). Sensitivity to Design: Early Understandings of the Link Between Agents and Order. *Proceedings of the National Academy of Sciences.*

Putnam, H. (1975). "The Meaning of Meaning," *Philosophical Papers, Vol. II : Mind, Language, and Reality*. Cambridge: Cambridge University Press.

Roskies, A. L. (2007). Are neuroimages like photographs of the brain? *Philosophy of Science, 74*, 860-872.

Roskies, A. L. (2008). Neuroimaging and inferential distance. *Neuroethics, 1*, 19-30.

Rottman, B. & Keil, F.C. (draft) Elaboration and Diversity: What Matters in Scientific Explanations.

Rozenblit, L. R., & Keil, F. C. (2002). The misunderstood limits of folk science: an illusion of explanatory depth. *Cognitive Science, 26*, 521–62.

Sloman, S. A., Love, B. C., & Ahn, W. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, *22*, 189-228

Smith, A. (1937). *The wealth of nations*. (Cannan edition ed.),, Modern Library, New York (1937).

Sodian & Kristen, S. (2010). *Theory of Mind: Towards A Theory Of Thinking*. Springer

Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience, 20*, 470-477.

Note: Some of the research described in this paper was supported by NIH grant R37- -HD023922 to Frank Keil