Research Statement

Laura Franklin-Hall
September 2015

I began my academic career in the sciences. I switched from neuroscience to philosophy out of a desire to explore, not just the functioning of the natural world, but also the nature of our inquiry into it. That inquiry can be divided into a collection of interconnected activities: experimentation, inference, categorization, theory construction, and explanation. Those aiming to understand one or more of these activities—both how they are carried out and what they ultimately achieve—face four significant tasks.

The first task is descriptive: to chronicle particular scientific episodes. What inferences have scientists made and in what contexts? What explanations and categorizations have they constructed or used? Here the philosopher's data are the practices of scientists as revealed in contemporary and historical writings, discussions with scientists, and direct observations of scientists at work.

The second task is theoretical: to uncover principles that describe—or perhaps regulate—the practices outlined in task one. Do scientists formulate those explanations that best unify phenomena or those that describe underlying causal mechanisms? Do they categorize entities and processes based on shared essential properties or in some other way? Sometimes scientists offer accounts of their own activities—and when they do, they are of philosophical interest—but often they do not.

The third task is evaluative: to ask whether the practices described in task one, and systematized in task two, are good ones. Since the working assumption of most philosophers of science is that scientists go about things, in large part, correctly, critiques are usually local and conceptually 'internal' to the scientific project as a whole. In particular, they proceed by pointing out ways that individual practices are in tension with other scientific results or commitments.

The fourth task is meta-evaluative: to probe the nature of the principles uncovered in the second task and subjected to scrutiny in the third. Are the correct principles objectively correct, such that compliance with them is required of all rational inquirers? Or might inquirers have done things very differently—e.g., might they have respected different explanatory, categorical, or inferential principles—without going wrong in any way?

My research to date has targeted three scientific activities—experimentation, classification, and explanation—largely, though not exclusively, in the context of biology. In each case I aim ultimately to contribute to all four tasks just outlined: the descriptive, theoretical, evaluative, and meta-evaluative. Though the details of my proposals vary, my work is thematically united in its attempt to grapple with a common finding of the descriptive project: the interestingly different forms that scientific practices like classification, explanation, and experimentation take across biology and other sciences. For instance, classifications can be either historical or synchronic, explanations either abstract or concrete, and experiments either hypothesis-driven or exploratory.

To make sense of this diversity, one option is to posit different categorical, explanatory, and experimental principles for different domains or contexts: perhaps explanations follow one recipe in ecology and another in molecular biology. A second, not incompatible idea is to account for diversity by reference to metaphysical differences between the systems under study: perhaps biological species, but not chemical elements, are classified historically because they are *individuals*, rather than *natural kinds*. A third option is to reject some actual practice as erroneous or non-ideal, making it less relevant for those formulating regulative principles: perhaps exploratory experiments are the mark of an immature science and good experiments are invariably hypothesis-driven.

The constructive part of my work pursues a fourth path: to articulate general principles according to which experimental, classificatory, and explanatory practice should take different forms depending both on the physical architectures of the systems under study, and on the capacities and interests of the inquirers who study them. This *flexible but unified* approach to the philosophy of science has many attractions. It is tolerant of scientific diversity. It offers resources with which to explain that diversity, not simply to describe it (which is arguably the result of simply positing different principles for different domains). And finally, the explanations it provides appeal to straightforward facts about physical architecture and inquirer capacities, rather than to comparatively obscure metaphysical claims.

Experimentation

Observation and experimentation are the foundation of the scientific project, and their judicious use among science's most characteristic features. My article "Exploratory Experiments" describes and evaluates a shift in experimental practice in biology that took place in the early years of this century. In this shift, experiments went from being almost exclusively hypothesis-directed to some taking a more 'exploratory' cast. I begin with a case study characterizing two experiments—one hypothesis-directed and one exploratory—that targeted broadly the same phenomenon: protein regulation over the cell cycle. Unlike the hypothesis-directed experiment, the exploratory one made use of a 'high-throughput' instrument able to make thousands of measurements simultaneously.

Some have doubted the worth of exploratory inquiries in recent biology and elsewhere. Are they not aimless 'fishing expeditions,' lacking the disciplined focus proper to science? The standard 'hypothetico-deductive' view of experimentation, on which experiments should test antecedently formulated hypotheses, would suggest as much. Even those philosophers who have been sympathetic to exploratory experiments usually restrict their legitimate role to the early years of a science, when its basic conceptual apparatus remains in flux.

My paper offers a more sympathetic account of the turn in experimental practice that my case study exemplifies. Though true theories may be their ultimate aim, the more immediate goal of many scientists is to identify difference-makers, and to do so as efficiently as possible. For those lacking 'high-throughput' instruments, the most efficient strategy for uncovering difference-makers is almost always by way of a hypothesis-directed experiment. But when such ‘high-throughput’ instruments *are* available—and when scientists have well-established background theories that place *some* constraints on what the potential difference-makers might be—exploratory experiments can offer the more efficient course.

Though I have not yet followed up on this early work, it provides a good example of the *flexible but unified* approach to the philosophy of science described above. While in this case, the variation in scientific practice I describe is accounted for by a variation in our capacities (here, our instruments), in other cases differences in the *physical architectures* under investigation play a major role. In my four completed papers on classification, described next, both factors come into play.

Classification

Scientific classifications—the neuroscientist’s division of cells into interneurons and astrocytes, the chemist’s of atoms into elements like carbon and oxygen, the psychologist’s of emotions into anger and fear—organize our universe’s constituents into categories and kinds. Such categories are important, among other reasons, because they provide the language in which generalizations, theories, and explanations are formulated.

My first paper on scientific classification looks closely at the nature of species among our planet’s dominant life form (judging by mass or census): bacteria. A customary view in the philosophy of biology is that species across the living world are population-level lineages, branches on the Tree of Life between speciation events. This historical, or ‘genealogical,’ approach to species is attractive, among other reasons, because it promises to deliver a unique and ‘non-essentialist’ classification—and one that does not require us to say just which properties of organisms are most important, as alternative ‘phenetic’ classifications arguably do.

How well does this view fit bacteria? Not well, I argue in “Bacteria, Sex and Systematics.” A first indication of difficulty is the fact that the ‘official’ definition of bacterial species—the one produced and used by working microbiologists—is (loosely speaking) phenetic, not genealogical. It identifies bacterial species as sets of organisms passing a test for genetic similarity and sharing one ‘diagnostic’ characteristic. Of course, this doesn’t show that bacterial species aren’t *really* branches on life’s tree; microbiologists, being the practically minded researchers that they are, may have formulated an easy-to-implement standard, without regard to the true nature of the underlying groupings.

My essay offers a different diagnosis, one that more directly vindicates the microbiologist’s distinctive approach. Due to an eccentricity of bacterial biology—their tendency to exchange genetic elements outside normal reproductive events and with partners of radically different kinds (Lateral Gene Transfer)—most bacteria have *many distinct histories*, not simply one. An organism’s ribosomal genes, cell wall genes, and antibiotic resistance genes may have originated from three different ancestors. To use this plural history to produce a genealogical classification of species, not to mention higher taxa, some have elevated a part of this complex history above the rest. For instance, ribosomal genes are sometimes singled out as defining a genealogical ‘backbone’. But why these genes and not others? Perhaps it is because they are genealogically ‘essential,’ but this will not be an attractive line to the many who aim to avoid such metaphysical commitment. On the other hand, if they are chosen on more practical grounds, it is not clear why the genealogical approach would be superior to the genetic standard microbiologists actually use. In the light of these complexities, my paper ends by recommending classificatory pluralism: the thesis that different domains of living things—or perhaps even each individual domain—may be properly classified in different manners, depending both on their intrinsic features and on the aims of the classifiers themselves.

The entry point for my second paper in this series—“Trashing Life’s Tree”—is a debate within biology concerning the theoretical worth of the Tree of Life, normally understood to be a recursively bifurcating graph representing the history of species lineages. Although it has been a model for life’s history since Darwin, prominent microbiologists, such as Ford Doolittle, have argued that because complete bacterial histories are *not* exclusively diverging, the Tree of Life model is literally false and should be discarded.

Yet the Tree of Life remains a scientifically popular representational scheme. Is its continued role well-motivated? My essay offers a limited defense of the special place of tree representations of life’s history even within microbiology. Among non-meiotic organisms like bacteria, *cellular* history is almost fully bifurcating, even if gene history is not. So, if we understand bacterial branches on the Tree of Life to represent a summary of the history of cellular division—rather than a full history of organisms or species—it will be closer to being veridical, if not perfectly so.

But why single out cellular history in this way? Is this another ‘essentialist’ move? Not necessarily, as there exists a more metaphysically innocent justification for this practice. Its starting-point is the fact that representations aren’t always better for being more complete. What should be included or left out depends on just which purposes a representation is supposed to serve. Though not all biological projects will be well-served by a map of cellular histories, many will be; cells in particular may have an important role to play in the dynamic modeling of bacterial evolution, given that the adaptive value of a gene almost invariably depends on the other genes with which it is packaged. I conclude that, in virtue of the cell’s importance in this regard, a bifurcating model of life’s history may well have a special place, not just in our understanding of genetically ‘well-behaved’ creatures like mammals, but even in our picture of the bacterial world.

My third paper in this series—“Natural Kinds as Categorical Bottlenecks”—develops the idea, gestured at above, that we tend to single out as special those representations that serve a large range of purposes. The focus of the paper is the problem of *natural kinds*. Often inspired by Plato’s famous remark that categories should ‘carve the world at its joints,’ philosophers—and to some degree, scientists too—have long distinguished between natural kinds, such as *gold*, from kinds that are artificial or gerrymandered, like *white things* or *objects within three miles of the Eiffel tower*.

Though there is relatively broad agreement about just which actual kinds are natural and which are not, there is no agreement on what property distinguishes natural kinds from nonnatural ones, nor about whether this difference is itself objective or 'real.' My essay addresses both of these questions. After proposing a framework for distinguishing realist and antirealist views in this context—itself a contentious issue—I argue that realist accounts of natural kinds fall prey to a serious epistemological objection. To avoid this objection, I propose understanding natural kinds in a more light-weight fashion as 'categorical bottlenecks': groupings corresponding to categories that well-serve our own purposes along with those of agents somewhat different from ourselves. This approach has a number of virtues: it permits the natural kinds to enjoy some level of objectivity, more than they do on competing antirealist accounts. Moreover, as it isn't always immediately obvious which kinds fit this standard, the account also makes sense of the customary claim that we have made progress over the history of inquiry in identifying the natural kinds. Finally, it avoids the epistemological problems associated with realism.

My final essay in this series, "Why are Some Kinds Historical and Others Not?", explores arguably the most substantial divide in classificatory strategies across the sciences: the fact that whereas some individuals and processes are classified *historically,* according to their relationships to past events (as when a bird's species is defined as a function of the population from which it descended), others are classified *synchronically*, according to shared characteristics (as when bacteria are grouped by their intrinsic genetic properties, or atoms by the number of protons within their nuclei). But why do scientists take a historical approach to some domains, and not to others?

Inasmuch as this topic has been tackled in the philosophy of science, it has been given a metaphysical diagnosis. According to the species-as-individuals thesis, for instance, biological species are not actually kinds—as Aristotle, for instance, conceived them to be—but are *individuals*, entities whose parts (the organisms) are necessarily spatiotemporally connected. On this view, the organisms that constitute a species will share a history in terms of which they should be grouped. In contrast, other scientific categories—each of the chemical elements, the physicist's fundamental particles, etc.—reflect *kinds* whose members need share no common origin but instead possess common intrinsic properties.

My view isn't that these suggestions are necessarily false, but that, from an explanatory point-of-view, they are superfluous; we need not appeal to metaphysical facts of this kind when accounting for why either species or other entities are categorized in the way that they are. My alternative proposal is that domains are—and should be—categorized historically just when the Probability of the Independent Emergence of Similar things (PIES) in that domain is extremely low; if PIES is high, then the domain should be categorized synchronically. This appears to correctly track scientific classificatory practice across the sciences; my key examples come from botany, microbiology, chemistry, and linguistics. And though space does not permit me to lay out the reasoning here, I contend that this approach is rational because, in honoring it, scientists are able to serve two important aims: (1) to develop categories in which it is possible to phrase strong generalizations; and (2) to categorize entities in a way that provides the best explanation of their shared features.

Explanation

An important dividend of the scientific project—some say the only one of intrinsic value—is explanation. It is via explanations that science permits us not only to predict and control events, but also to *understand* the universe and our place within it. My work on this topic aims to articulate principles that describe and regulate scientific explanatory practice, as well as to assess the objectivity of those principles.

The feature of explanatory practice on which my efforts have focused to date is what I call explanatory 'sparseness': the fact that perfectly acceptable—maybe even optimal—explanations often omit mention of what can appear, from a physical point of view, to be critical details. Along the *horizontal* dimension, explanations often do not state complete *necessary conditions* for the events or regularities for which they account. Thus, an explanation of a neuron's firing may not mention the fact that its membrane is intact, though that is indeed necessary for the occurrence of the target event. And along the *vertical* dimension, explanations often cite only 'high-level' or coarse-grained states of affairs; a concentration of some protein, for example, rather than a list of precise molecular locations.

Why does our explanatory practice have this character? One reply, most apt when accounting for the vertical dimension of explanatory sparseness, is metaphysical: perhaps there exist 'top-down causes' or strongly 'emergent' properties that high-level explanations quite properly track. Another reply is purely practical: that sparseness results from our lack of knowledge of the objectively relevant physical details, or from difficulties presenting those details even when known. On this view, optimal explanations would have a rather different character than those that biologists and psychologists (for instance) actually articulate.

The metaphysical suggestion founders, I believe, on the considerable evidence that our universe is through and through a physical one, governed exclusively by 'low-level' physical laws acting on physical systems. The practical suggestion has more to be said in its favor, and it would not be easy to conclusively show that it is mistaken. After all, there is no doubt that we *do* often lack the knowledge required to articulate comprehensive physical accounts. Still, philosophers—and, to a degree, scientists themselves—often express the view that scientific explanations are frequently *better* for being abstract, suggesting that the sparseness of actual explanations might have a more principled, explanatory motivation.

Yet it has proven difficult to substantiate this hunch, particularly for those who take an exclusively causal (rather than 'unificationist') approach to explanation. For instance, in my two critical papers on this topic—"New Mechanistic Explanation and the Need for Explanatory Constraints" and "High-Level Explanation and the Interventionist's 'Variables Problem'"—I argue that the advocates of the two most influential accounts of explanation in the philosophy of biology have not yet succeeded in either characterizing the correct 'level' at which explanations should be offered, nor have they successfully explained just why higher-level or 'sparse' explanations are more desirable than lower-level accounts. In particular, I suggest that the interventionist's notion of a 'variable' and the new mechanist's understanding of a 'part' need further buttressing before the accounts in which they feature might be able to 'save the explanatory phenomena.'

The problems that I press against these two views are hard ones, and the aim of my critical work is not to condemn these projects. Instead, my hope is to draw attention to the very interesting unresolved issues that lie at their hearts. In fact, I focus on these targets because I agree with them on many important matters, and in particular about the fact (1) that scientific explanation in the high-level sciences is fundamentally causal in character; and (2) that high-level explanations take the form that they do on principled explanatory grounds (not merely practical ones).

Yet rather than try to patch up these approaches, my constructive explanatory project has been to formulate an alternative view: the Causal Economy Account. On this theory—which is described, with different emphases, in "Explaining Causal Selection with Explanatory Causal Economy" and "The Causal Economy Account of Scientific Explanation"—an event's correct explanation will cite that subset of its causal influences which is (metaphorically speaking) maximally 'economical.' An economical account says as little as possible about an event's causal run-up (thus being descriptively 'cheaper') while mentioning just those causal factors that make the target event, to a substantial degree, *bound to happen* (thus 'delivering more'). On this view, any account that maximizes economy—and there may be more than one of these—constitutes a complete and correct explanation.

I characterize both 'cost' and 'delivery' in a possible worlds framework. Very briefly, an explanation costs more to the degree to which it contains more descriptive content, ruling out more ways that the world might be. And an explanation 'delivers more' to the degree that the event it explains would have occurred in a larger number of nearby worlds in which we have fixed, via a small miracle, the occurrence of the causal influences that it cites.

As a rule, more elaborate descriptions of an event's causal influences will cite facts that more fully necessitate its occurrence. But some systems include what I call 'sweet spots,' physical architectures in virtue of which just a few causal factors have a disproportionate influence on an event's occurrence. What are customarily called 'biological signals,' such the vervet monkey's 'leopard' alarm call or the cytokine molecules that recruits immune cells, appear to this character: they are rather cheap to characterize, yet they have very systematic and robust effects on the rest of the system. In particular, given that they are present, lots of others things might have been different, yet the events they are used to explain would still have taken place. I submit that it is partly due to the existence of such sweet spots in the architecture of our universe that high-level explanations—and the high-level sciences themselves—are possible.

One advantage of the Causal Economy view is that it traces the two types of explanatory sparseness noted above—'vertical' (high-level) and 'horizontal' (exclusion of background conditions)—to the same root, and in this way provides a pleasingly unified picture. Yet, at the same time, in the spirit of a 'flexible but unified' approach, the account makes sense of an important kind of diversity in our explanatory practice: the fact that a single phenomenon can sometimes be explained in either a low-level, mechanistic way, or in more abstract terms. This diversity is permitted—and perhaps even expected—because explanations making different trade-offs can be *equally* economical: some will cost little and deliver little, while others will cost quite a lot but deliver a great deal.

My final completed paper on explanation addresses, as its title indicates, "The Meta-Explanatory Question": are correct explanations objectively correct, or is explanation ultimately relative to our concepts, aims, or capacities? While in ethics it has been customary to ask a parallel question about the objectivity of the normative facts and theories, philosophers of explanation have been less attentive to the meta-explanatory question. The central aspiration of this short paper (written for a venue that limits contributions to 5000 words) is to draw attention to it.

The paper begins by formulating the question of explanatory realism and by offering a framework for distinguishing realist views from antirealist ones. Given that realism is usually the default position for any discourse we take seriously, the paper moves on to sketch three distinct paths to antirealism. The best-known of these stems from the deeply pragmatic view that there are no general explanatory principles of the kind philosophers usually try to articulate, holding instead that an explanation's correctness always depends on the particular context in which it is offered or received. A second path to antirealism, which may be more popular though has not been as clearly articulated, is 'functionalist.' Unlike the pragmatist, the functionalist endorses the search for general explanatory principles, but insists that what makes certain principles the right ones is that they best serve our practical aims (for example, those of predicting and controlling our environment). The third path to antirealism is based on an epistemological objection to realism (and resembles the challenge I pose to realism about natural kinds in "Natural Kinds as Categorical Bottlenecks"). The problem is that we seem to have no good reason to believe—and possibly some positive reason to disbelieve—that our scientific practices, which form the starting point for our theorizing about explanatory principles, track the objectively correct principles. If one is impressed by this problem, then it may seem more reasonable to think that the right explanatory principles depend on our aims or ways of seeing the world.

I myself am sympathetic to this last path to antirealism, and in future work I aim to make the case for it more strongly than was possible in the short paper just described. This antirealism might seem surprising, given that I offer an account of causal explanation that aspires to be relatively general, applying at least across biology, if not beyond it. Why work to construct such an account if explanation is not ultimately objective? Though they may appear in tension, these projects are compatible. While I suspect that the rightness of the right explanatory principles is determined by something about our concepts and capacities, these features may be deep enough in our nature so as to mean that there is just one set of explanatory principles that applies across the sciences—that is, across just those sciences that are carried out by human inquirers.