Ifs, Ands, and Buts: Defeasibility Management across Disciplines

Elijah Millgram

Department of Philosophy
University of Utah
Salt Lake City UT 84112

elijah.millgram@gmail.com

September 29, 2019

1 Defeasibility and Specialization

The Land of OR is a recent children's book that aims to instruct very young readers in decision-making (Mullaly and Allen 2015). For instance, it tells them to assemble a large menu of options (which are personified as blue or pink characters with antlers), and then narrow it down. Along the way, it introduces the "Yabbut," another anthropomorphic character (purple and green this time) that you're instructed to ignore: "The Yabbut you see, it has only one use; To always be near and provide an excuse." I should say that everything this book tells its deciders-in-training is 100% platitudinous; the parents doing the actual bedtime reading are not in for any surprises.

Nonetheless, dismissing those "Yeah, buts" is not just bad advice, but logically confused. Almost all inferences are *defeasible*, meaning that even though a step you're about to take in your reasoning is fine as is, you could add further considerations to it (say, if additional information turned up) that would *defeat* it, and here's a toy example, lifted from a popular column about the foibles of business executives (Suter 2000, 10–18). You're about to be promoted, and when you are, you'll get to redecorate your office. It's important for the interior decorating to send the right message, and also for a newly-promoted manager to act quickly and decisively; so you should be studying furniture catalogs *now.* Well, here's a defeater for that inference: Yeah, *but* spending your evenings in the office studying furniture catalogs is a great way to ensure that the promotion goes to someone else. You shouldn't actually draw that conclusion: you shouldn't be staying late with those furniture catalogs. When reasoning is deductive, the truth of the premises guarantees the truth of the conclusion; but occasions for deductive reasoning are rare, and most of the reasoning you or anyone will do is (and here's another term for "defeasible," used in fields like AI) *nonmonotonic.* (For overviews, see Hlobil 2018, Horty 2012, and Reutlinger et al. 2017.) The confusion is on display in that children's book I started with: a little further on, the pupil is warned to watch out for unexpected outcomes—but a very typical instance of a defeater (of a "yeah, but") *is* noticing that this time, there's going to be an unwanted consequence. (The unintended consequence is a reason to take back that initial conclusion.) You're told to ignore "yeah, buts," and *also* told not to ignore upshots of the pending decision—but these are (much although not all of the time) the very same thing.

Defeasibility is pervasive: it's not just a feature of decision making. When you float an argument for a factual or explanatory claim, it also almost always comes with *ceteris paribus* ('other things equal') clauses, implicitly or explictly. For instance, you might think it stands to reason that propagule pressure—the frequency of occasions on which a nonnative species is introduced into a new environment, and the number of organisms per occasion—ought to go a long way toward explaining when nonnative species successfully establish themselves. After all, each nonnative organism has some nonzero probability of being the foothold for that successful invasion, and sometimes the extra introduction effort is seen to be correlated with greater success. But there are various potential defeaters for this bit of reasoning, too, and we'll just point to the start of the list. It doesn't matter how many camels we drop into the Pacific, because practically speaking, each camel has a zero probability of surviving, and there isn't going be a swimming camel colony. The conclusion won't follow if the probability of success of an introduction isn't independent of other introductions: perhaps not if intraspecific competition within a sufficiently large population undercuts successful interspecific competition, and perhaps not if a too-large-too-quickly population gets the nonnative species noticed by the ecosystem—'noticed,' if predators or parasites adapt to the novel prey or host, or literally noticed, if humans are annoyed enough to attempt eradication. And the supporting evidence is undermined if the invasive "species were successful... not because they were introduced in large numbers, but rather species were introduced in high numbers because the initial releases were perceived to be successful and useful" (Moulton et al., 619).

When you lock down a claim or make a point—even if what you're doing is an experiment, in a lab—you can represent how that point got made by spelling it out as an argument. If you're thinking about it that way, and I'll get back to this issue in due course, any experiment you do, any study you perform, is going to involve defeasible reasoning.

Now, the bad news is that defeasibility isn't well-understood. Maybe the deepest logical feature of the phenomenon is that potential defeaters don't run out; for any nondeductive inference, you can always think of further considerations such that, if they were in play, they would be good reasons not to draw its conclusion. (If you're not sure, start a list of reasons someone might have to balk at that argument for reading office furniture catalogs: you'll discover that no matter how long the list is, you can always think of another one.) We're not very good at knowing what to do with this sort of open-endedness. Most of the advice out there amounts to, 'Look around and try to understand what might go wrong with your train of thought.' But since there are always more things to check, how do you tell when you've checked *enough*?

So the general problem is hard enough already. But it seems to become even harder when we think about how it's inflected by the need to cooperate across different areas of specialized expertise. It's all very well to tell people to work at understanding what might go wrong with a train of thought. But you depend on results and on tools that are provided by other specialists who are differently trained than you are: their training probably took up of a decade. In most cases you're just not in a position to know what they do, or to know your way around their field, or, consequently, to know what might go wrong. (For a medium-elaborate illustration, see Hardwig 1985; see also Jeschke et al. 2019 and Nguyen 2018.) The default advice—just *understand*—is unrealistic.

Let's pause for a moment to spell out the problem a bit. Part of the training people get as students is normally in field-specific vocabulary, notation, and some repertoire of (often) mathematical tools; if you didn't get that training, you probably can't so much as read their literature. Apprentices in a field get inducted into a system of discipline-specific standards; if you're not in that field, you haven't internalized those standards, and so you're not in a position to do quality control, even on what you *can* read. Students are trained on procedures, and a lot of that is tacit knowledge-how (it's never fully spelled out). And one very important part of that discipline-specific knowledge-how is knowing what to watch out for, say, when you design and run and interpret an experiment: which potential defeaters you have to be aware of and check. Knowing what to watch out for will mostly be something only specialized insiders are equipped for, not least because explaining and training people to this sort of competence presupposes that they already have mastered the field-specific representational vocabulary and standards.

So now consider how this plays out, on the assumption that a particular experiment requires the cooperation of, for now, just two differently specialized researchers. One of them isn't in a position to figure out what might go wrong, because he's helping himself to the other one's resources—information or procedures or equipment, for instance—that he's not in a position to really understand. But the other isn't in a position to figure out what might go wrong for his collaborator (or client, if he's a core facilities provider), because he doesn't really understand what use is being made of that information (and so on). The symmetries of this sort of situation suggest that you can't rely on *anyone* to catch the defeaters for these sorts of arguments.

You might think I'm overstating the difficulties. After all, when you have to work with people in other fields, you develop interactional expertise (Gorman 2010); you do catch problems and get things to work; you even get good at it. There are various things we routinely do, in daily life even, to improve the way these sorts of issues play out. But remember that anticipating *some* problems isn't the same thing as being able to be confident that you've identified (*all*) the important ones. Because we don't know about what we haven't noticed, we tend to overestimate how good we are at this.

And perhaps there's less in the way of preemptive coping than you'd think. Every now and again, I conduct anecdotal surveys. I explain the problem—most recently, to software engineers, testing staff for a pharmaceutical company, marketing consultants, investment portfolio managers and stock analysts—and ask what procedures they've put in place to deal with it. Dishearteningly, the usual answer is: we haven't. For instance, I was told by fund managers that they diversify, so that overlooking one or another defeater wouldn't turn into a disaster, but also that it's normal to start looking into problems *after* a firm begins to lose market share. Out in the nonacademic world, there seems to be surprisingly little in the way of attempts to *anticipate* this sort of problem.

At this point in the exposition, it's reasonable to expect an extended illustration. However, the pivotal feature of the issue we're considering requires us to postpone that example; come the tail end of the paper, we'll conclude with one that is somewhat indirect, and let me explain why.

What we're worried about is that the need for disciplinary competences which one does not have prevents one from understanding what the potential defeaters of an inference are. But if the illustration is of *that*, then the reader won't understand it, and then how can it help keep a reader on board? To be sure, there are a couple of plausible workarounds. First, we could give a toy example of the phenomenon we are discussing; the reader *would* understand all of its moving parts, and be asked to imagine scaling it up to something he *doesn't* understand. In this case, the drawback is that we're not really getting an illustration of the thing we wanted, namely, something you can't understand; and so it would be quite reasonable to balk at claims I'm making, to the effect that cross-disciplinary defeasibility can't be handled just by asking people to understand more. Alternatively, we could introduce a very elaborate example, and spend the time needed to explain all of its parts to the reader. Here the drawback is that, as in real life, the costs of fully understanding material drawn from different fields or subfields are too steep to ask readers to incur. (But there are other options: for a different sort of workaround, see Millgram 2015, 50n49.)

So the illustration I plan to conclude with will compromise, and do a bit of both. Over the course of the discussion, I will put in place a relatively complex case study, one which I hope will be clear enough to be persuasive, even if it is much less elaborate than a completely realistic treatment would be. To anticipate, the illustration will turn out to be the very idea that underlies the approach I'm about to introduce.

2 The Hierarchy-of-Hypotheses Approach

The way I've framed the discussion suggests that the special case of defeasibility management across disciplines has to be a hard problem, much harder, in fact, than the general version we started with. That seems to follow from how it is that no one is in a position to understand what might go wrong with all those different arguments that make science (and not just science) go around. But here's where we're going: maybe the special case is *easier* than the general form of the problem. To make that plausible, I'm going to take up a recent methodological discussion in invasion biology, that is, the biology of invasive species, centered on the so-called Hierarchy-of-Hypotheses (or HoH) approach (Heger and Jeschke 2018).

During the second Gulf War, Donald Rumsfeld famously said that you go to war with the army you have, not the army you wish you had. And something like that goes for research also. When you are going to test a high-level, relatively general hypothesis, you go about it with the laboratory, or field station you have (not the one you wish you had); you draw on the training you have, and so on. That's true across the board, and the special case I'm most interested in is the one where tests of an hypothesis are going to be conducted by people with different disciplinary backgrounds. They'll go at it with the training they have, and the tools their disciplines provide, not the training they would have gotten and the tools they would have had, had they been in a different subfield. Let me emphasize that it is a special case: even within subfields, researchers are trained on and have experience with different approaches to problems; different laboratories are practiced in applying different techniques. Nonetheless, confining ourselves to the special case will keep the shape both of the problem, and the solution to it that I'll shortly broach, in clearer focus.

You start with an overarching hypothesis—say, 'enemy release,' the suggestion that invasions work out when the new environment doesn't have the predators or parasites of the species' home range—but then you have to make it more concrete, and this happens in stages: you confine your attention to a working hypothesis, maybe to do with, borrowing an example from Jeschke and Heger 2018, 16, 31, "rates of attack by natural enemies" and "invasiveness of exotic plants"; finally you turn it into an operationalized hypothesis: you'll compare invasive and native buried seeds, to see which are more vulnerable to fungi. The way you end up making that general, shared hypothesis more concrete will be driven in large part by the research program to which you committed, which will in turn be constrained in the first place by your disciplinary specialization. ('In the first place': once again, the different investments which have been made by different researchers who share a single subfield can be equally important.) The stage-by-stage multiple concretizations of an hypothesis will form a hierarchy, often representable as a tree.

At lower levels, the hypotheses will look different—even *very* different—

from one another. The differences can go all the way down to what versions of the concepts are in play; as Jerome Ravetz pointed out long ago (1979, ch. 6), what is intuitively the very same concept will be differently introduced, defined and tied to measurement techniques in adjacent fields. When you ask differently specialized researchers to investigate one and the same hypothesis, you will often enough see it firmed up into what look like qualitatively different questions. (That such different approaches are being pursued isn't always obvious; for discussion of a case in which two research traditions talked past one another for a generation, see Tabery 2014.)

Let's anticipate a sticking point: maybe what *look* like qualitatively different questions just *are* different questions, in which case it would be a mistake to see those different approaches as homing in on *one* hypothesis (compare Griesemer 2018, 25). We have recently seen complaints of this kind—that there is only the illusion of a shared research program—in a nearby field; the objection is that it is only by equivocating on the concept *biodiversity* that we think that different researchers are investigating *one* phenomenon (Santana 2014, Morar et al. 2015, Santana 2018). We should, the complaints conclude, eliminate the concept that is causing all the confusion, and that proposal has come to be called "biodiversity eliminativism".

Putting that worry to one side for the moment, the upshot is that tests of such an hypothesis will be very differently inflected. Again, any such test amounts to—that is, can be represented as—an argument, so let's keep on thinking of them that way. And think of the potential defeaters for an argument as arranged into an open-ended list, prioritized so as to put the more urgent, more important potential defeaters closer to the front. Then, by and large, different arguments will turn out to have their own distinctive (open-ended) lists of defeaters. Although there can be overlap between these lists—a potential defeater will appear on more than one of them—when you're looking at two such lists that share an entry, typically you'll find the defeater appearing higher up or lower down on one of them.

Now we said that different specializations concretize overarching hypotheses into differently operationalized hypotheses. Because different specializations bring to bear their own standards for constructing and assessing arguments, that means that these different tests will amount to substantively (and often formally) different arguments.

Even waiving the worry we introduced a moment ago, this seems to have a large downside, namely, an intractable assessment problem. After all, the tests are so different; how do you compare them? The researchers spearheading the HoH approach have opted for an extremely stripped-down basis for commensuration: you code the overall direction of support of an experiment or study, in one of three ways; you sum up those thumbs up and thumbs down and thumbs sideways; and optionally, you then tag the nodes of the graph of the hierarchy of hypotheses, either with the overall direction of support, or with the distribution of studies into the supporting, unsupporting, or undecided categories. In a recent anthology, the procedure indeed attracted complaints, which in one way or another had to do with lost nuance, and the risks involved in handling information so roughly (Jeschke and Heger 2018, esp. ch. 3; cf. also 95–97).

But there's also a very large upside. It follows from the points we've already made that the different arguments for and against an overarching hypothesis, when they're developed within different specializations, will normally travel along with different implicit (again, open-ended) lists of defeaters. And this means that if there are *many* differently concretized tests of an hypothesis generated within different fields, an unnoticed potential defeater for one of them will not necessarily be—even, will often not be—a high-priority defeater for others. Because the methodologies are different, the implicit and explicit arguments are also different. . . and consequently, the defeasibility conditions for those arguments will be different. And this means that the HoH approach allows you to test an overarching hypothesis, in a way that *permissibly ignores* defeaters, in virtue of checking the overall direction of support: the one that preponderates in the class of tests of the concretized subordinate hypotheses.

This way of assessing support for an overarching hypothesis doesn't require you to *understand* the nuts and bolts of each study. Recall that this was what we found ourselves unable to do, when experiments deploy resources that cross disciplinary lines. And so what seems like an obstacle is actually an advantage. We had thought we needed the too-expensive-to-be-possible understanding in order to locate defeaters for our inferences. But when there is enough disciplinary and methodological variation, we can expect the different defeaters of differently constructed inferences to wash out—in which case, we don't actually have to identify them. Thus the very crudeness of the assessment procedure, which prompted those complaints, turns out to be a positive virtue: if the studies were directly comparable, and you compared them at a higher level of granularity, you wouldn't be able to treat the defeaters as noise.

To recap, to someone interested in interdisciplinary defeasibility management, the Hierarchy-of-Hypotheses approach is promising for two related reasons. It implicitly acknowledges the ways that, when you ask differently specialized researchers to investigate a given overarching hypothesis, you will see that hypothesis firmed up into what seem like qualitatively very different concrete questions. And that very diversity of reframings allows us to sidestep the impossible task posed by the usual advice we're given for managing defeaters ("just *understand*"); what is a defeater for the argument built into one way of answering one concretized question won't normally be a defeater for all or even most of the others.

3 Potential Defeaters for the Hierarchy-of-Hypotheses Approach

How optimistic should we be? Is the way forward clear? We have already introduced one objection, and I'm going to survey a handful of further problems that have to be surmounted, if the promise of the approach is to be made good. The response I'll suggest, by way of wrapping up the discussion, is to go meta: to turn, not exactly the nuts and bolts, but the spirit of the HoH approach on itself.

To anticipate, the HoH approach is one of a number of methodological approaches that share a motivating insight. Because they are couched in very different intellectual vocabularies, the objections to any one of them are unlikely to be objections to the others. And that in turn suggests that, if these approaches have a largely successful track record, the underlying insight, and thus the HoH approach itself, is likely to be on the right track.

Turning to a review of those problems, that is, objections to the train of thought that endorsed the HoH approach as a way of managing defeasibility: First, and apropos those similar approaches, in the emerging literature on robustness analysis, people worry about the so-called independence problem. (We will introduce that discussion shortly.) Perhaps there is more overlap between those different experimental methods than you think; perhaps what is a defeater for one of them is a defeater for many, or even all of them. So when your approach is to survey the applications of different experimental methods, don't you need to check that the methods are independent—for present purposes, that their defeaters really don't overlap, anyway very much?

But now, surely the only way to find those overlaps and verify that they are safe is to understand (and even understand deeply) how those tests in the several disciplines are firmed up and conducted. And that very demand—*just understand*—was the task we thought was too hard to impose, or live up to. (After all, unless you are competent in the many relevant fields, you cannot tell whether the problem is real.) So, the objection has it, a technique that was supposed to substitute for a limited resource (namely, understanding) turns out to presuppose that the resource is unlimited.

Second in our short list of problems, let's go back to the sticking point we briefly worried about earlier, that the process of converting an overarching hypothesis into more concrete and testable versions of it just gives us *different* problems. We can first parry the objection and then sharpen it up. That process is analogous to the process of firming up a thinly described goal into an objective that is concrete enough to anchor a search for means to it; in the so-called specification of ends, a single abstractly rendered goal can, depending on what the background constraints are, be specified as any one of many substantively different targets (Millgram 2008; cf. Jeschke and Heger 2018, 15). We would lose the flexibility we need in figuring out what to do by insisting that the more concrete goals are, not different specifications of the more abstract one but, rather, just *different*; we have to be more relaxed in how we think about objectives.

Let's just concede that we would likewise lose our ability to wash out defeaters by insisting that distinct tests or procedures mean distinct target concepts, and concede further that this is a similarly good reason to be more relaxed about what counts as a tightened up version of an overarching hypothesis. Going back to the complaint we gave as our example, namely, that the biodiversity literature operationalizes its central concept in too many different ways, we're proposing to allow that imprecision of this kind is actually a valuable resource, one that is worth conserving. It's a *good* thing that biodiversity is sometimes firmed up into species richness—the number of species found at a given place—sometimes into phenotypic diversity, and sometimes other ways.

But still, surely, the problems being addressed in different disciplines may nonetheless only *seem* to be the same problem from a distance. As before, wouldn't the only way to tell involve understanding the more concrete, intermediate hypotheses, and the methods used for testing them, in depth? And as before, wasn't that the task we thought too hard to impose, or live up to? And indeed, continuing with the example, an outsider like myself can't tell who is right about biodiversity eliminativism—remember, the methodological suggestion that we remove biodiversity from our repertoire of scientific concepts, and couch our conclusions and policy proposals in terms of species richness, phenotypic diversity and so on, *directly*.

Third, even if the HoH approach works as advertised, and allows us to disregard the *ceteris paribus* clauses in the arguments for a given hypothesis, especially in fields such as invasion biology, an hypothesis, once accepted, is meant to be applied. An argument that has such a confirmed hypothesis as one of its premises and a policy recommendation for its conclusion will also be defeasible, and its users will nonetheless need to remain alert for the indefinitely many defeaters pertinent to the proposed application. The enemy release hypothesis suggests managing invasive species by importing their enemies; as an old film, *Cane Toads*, amusingly reminds us, there are many ways such a policy can prove to be ill-considered, just for instance, when the imported enemy itself undergoes enemy release (Lewis 1988). So in practice, only *some* of the defeasibility that we need to cope with seems susceptible to HoH-based management. Relatedly, as Elliott-Graves 2016 points out, even if your hypotheses capture something real, that's a long way from being able to generate specific and usable predictions.

Fourth, the method is limited to those cases where workers in different fields are testing the same overarching hypothesis. Only some cases of cross-disciplinary cooperation take this form; often, instead, one expert works on his part of a problem that spans several, even many, disciplines; another works on a different part of the problem, and so on. In such cases, they're not producing different kinds of argument for the same conclusion; and then we're not in a position to aggregate their work, in a form that will reasonably allow us to presume that the defeaters for their arguments are washing out.

Fifth, and last for now, what assessment results from the procedure we described above is evidently sensitive to how studies are individuated. (E.g., if we are counting publications, and what is really one study has been published three times, we will end up with skewed outcomes.) The home domain of the HoH approach is invasion biology, where it helps to eyeball the locations and principal investigators in order to distinguish one node in the hierarchy from another. But the ambit of the approach is potentially much broader; when we are including, to take an extreme instance, the arguments of philosophers, how are we supposed to count up those arguments? Wouldn't developing and applying criteria for individuating studies, experiments, and more generally arguments require intimate knowledge of all of their disciplinary homes, that is, knowledge that no one could have?

Think of these as the beginnings of a list of defeaters for our argument in favor of the HoH methodology; as we observed early on, we should assume that list to be open ended. And at this point, you may be wondering whether we have made any real headway; after all, a pattern seems to be emerging. For all but the fourth of the problems we have just introduced, vetting for the problem in typical instances seems to require hands-on, nuts-and-bolts understanding. That understanding will often require disciplinary expertise in more than one discipline. But recall that this was the limiting resource. So let's turn to a reason to think, despite those concerns, that we should continue to develop the HoH approach.

4 Going Meta: Starting a Hierarchy of Methodological Hypotheses

The HoH approach is not the only methodological attempt to bypass the difficulties of defeater-by-defeater management of defeasible inference. It shares an overarching insight with (at least) two related approaches, namely, that what we are after is a way of assessing the overall confidence we should have in our conclusions, by seeing if glitches can be treated as noise.

Readers with a background in philosophy of science are likely already to have been reminded of Richard Levins. Rehearsing his well-known methodological summary,

> we attempt to treat the same problem with several alternative models each with different simplifications but with a common biological assumption. Then, if these models, despite their different assumptions, lead to similar results we have what we can call a robust theorem which is relatively free of the details of the model. Hence our truth is the intersection of independent lies. (1966, 423)

In the wake of the proposal, there has been ongoing attention to robustness analysis, amounting at this point to a program (Weisberg 2006)

And then also there has been a recent attempt to make sense of an hypothesis being robustly confirmed in terms of invariance over multiple means of detection. (Schupback 2018; however, I need to add a parenthetical caveat: I'm going to strongarm Schupbach's treatment to suit our own purposes. The paper has two laps, the first of which presents that initial idea, and suggests that Levins's proposal is a special case, where his various models count as different 'means of detection'. It's reasonable to construe this as a way of licensing the sort of confidence in an hypothesis that would allow one to dismiss complaints about one or another means of detection as noise. But the second lap advances a proposal on which a robustness analysis for an explanatory hypothesis proceeds by checking one after another means of detection, such that each additional check removes a competing explanation for the hypothesis. The proposal takes off in a direction that's no longer in the spirit of the idea we're interested in; to dismiss competing explanations only after you've checked them is not to dismiss them as noise without checking them, on the basis of the overall confidence that has accrued to one's hypothesis.)

That allows for a little compare and contrast. First of all, each of these three approaches has its home in a different disciplinary specialization. The HoH approach grows out of work in invasion biology; Levins came to his ideas from population biology; Schupbach's treatment has its home in formal epistemology.

As our argument up to this point would lead us to expect, the conceptual vocabulary that each of them deploys is different. We have just seen some of the equipment in the HoH approach's toolkit. Both Levins and Weisberg are focused on causal structure which can be represented mathematically; the discussion is conducted in terms of models. The formal-epistemology approach represents its claims in the language of Bayesian probability.

Again as we would expect, the forms that techniques for assessing hypotheses take in the three approaches differ. As we have seen, the Hierarchy-of-Hypotheses approach captures the overall direction of a number of experiments or other studies. Levins-Weisberg robustness analysis stress-tests causal models, with the aim of setting expectations about sensitivity to potential defeaters, and here's a very quick example (not one of their own). About fifty years back, a widely publicized book called *The Limits to Growth* (Meadows et al. 1974) argued that if the global economy continued to grow exponentially inside the closed system constituted by our planet, it would entail a large-scale economic collapse. The conclusion was held to be robust, because it turned on qualitative features of exponential functions; adjusting the parameters of their economic model one way and another made only marginal differences to its behavior. Finally, Schupbach has constructed a Bayesian index of explanatory power, and he invokes it to explicate the conditions he wants to impose on successful steps of a robustness analysis.

And once more as we would expect, arguments conducted within the different approaches come with different potential defeaters. I'll give just one example, an objection to the robustness-analysis approach that was not an issue for the HoH approach. (But as before, we could generate an indefinitely long list of possible defeaters for each approach.) Any model provides a skeletal representation of the world, in which most features of the phenomenon being studied, the circumstances in which it is embedded, and the interconnections between them simply do not appear; that sort of simplification is the very point of having a model. But that means that phenomena depending on the very density of features, connections and so on in the real world can easily fail to be captured by a family of models.

The potential defeater for Levins's argument for his way of testing robustness reappears as a candidate defeater for arguments conducted using his technique. Going back to that mid-seventies example, the simplified models of the global economy that gave the Club of Rome's book its title left out real estate markets, as well as markets for virtual commodities and the like. Sky-high housing costs in Palo Alto mean that ever more money can change hands without burning through any more in the way of oil, farmland, iron ore and so on—or anyway, allowing that server farms consume power, *much* more. The market in web pages makes the economy larger, even if fewer trees are being cut down. That is, the models miss (almost inevitably, in retrospect) ways in which economic growth is only loosely linked to what an economy burns through in the tangible world.

Recapping, the benefits of having the investigation of a high-level hypothesis conducted by researchers in different disciplines were (in part) mediated by their typically very different intellectual vocabularies. Arguments formulated in such different vocabularies will almost inevitably be *different* arguments, and consequently be vulnerable to different defeaters. Lo and behold, we can see this effect play out in the approaches to defeasibility management that we're now putting side by side: they stand in roughly the relation to that overarching insight that subsidiary hypotheses, in the HoH approach, stand to their overarching hypothesis.

Recall that, early on, I began to describe experiments, surveys and so on as *arguments* (allowing that the argumentation might be only implicit). I have found that way of talking to provoke replies from academics trained in other fields, along the lines of: an experiment isn't an *argument*—it's an *experiment*. We can now see that those replies reflect the way that different fields have different proprietary intellectual toolsets. The philosophers are used to representing things as arguments—it's their medium, their theoretical paint—and so when a philosopher addresses a problem, or tries to make a point, it will normally be rendered in that medium, and often seem to be about argumentation. Whereas Levins-Weisberg-style robustness analysis is conducted in a rather different medium, that of models. Formal epistemologists in turn have their own favored intellectual medium: they support their claims by, to put it a little baldly, deriving equations about probabilities that involve Bayes' Theorem. As expected, those different intellectual vocabularies generate different modes of argumentation. And that in turn means that the pitfalls to which one of these styles of argument is vulnerable are likely not to be encountered by the others.

Recall that we found ourselves worrying that a piece of our exposition would have to go missing. Because the problem we are investigating turns on obstacles to understanding, it seemed that we would be unable to provide a concrete but understandable illustration, either of the problem or its resolution. After all, if you understand it, it doesn't illustrate the point that there are things you can't understand; but if you don't understand it, what's the point of the illustration? However, we have now worked our way around to an accessible, albeit somewhat toy, example, namely, how to vet the overarching insight that we have been discussing, bearing in mind that we have just looked over three or so variants on it.

Remember that we began to survey the defeasibility conditions for the HoH methodology, and it started to look like vetting it would require the sort of multidisciplinary competences that no one could realistically be expected to have. But if we find that the variants on our short menu can mostly be gotten to work in the field, our confidence in that overarching idea should plausibly increase—as the spirit of HoH approach would indeed have it. It should increase even if we do not think we have identified all of the potential defeaters either for arguments in favor of the approach, or for the arguments underlying applications of it. It should increase even if no one could have the competences in all of the disciplines necessary even to articulate, much less provide, a principled resolution of the defeasibility conditions for those arguments.

So what should our take on the HoH approach be? Even though our list of potential defeaters for it was short, we know very well that it can be extended indefinitely. That suggests that the theoretically most satisfying support for the HoH approach would itself take the form of an HoH assessment—adapted, of course, to suit the nonexperimental content of its overarching insight. As our rough survey of ongoing discussions of robustness indicates, it would be premature to conduct that assessment now; we have described only three variants on the insight underlying the approach, which is a far cry from the dozens or even hundreds of studies envisioned in a satisfying application of the HoH methodology. But as our confidence grows in the overarching insight, that confidence should be transmitted to this way of concretizing it, and the methodology will come to seem more promising. And it does already seem to have perhaps put its finger on something very important.

5 Acknowledgements

I'm grateful to Chrisoula Andreou, Melinda Fagan, Tina Heger, Arnon Levy, C. Thi Nguyen, Carlos Santana and Jonah Schupbach for comments on an earlier draft, to feedback at the Workshop on Research Synthesis hosted by the VW-Stiftung, and at the Intermountain Philosophy Conference. Thanks to the Hebrew University for a Lady Davis Fellowship, and to the University of Utah for support through a Sterling M. McMurrin Esteemed Faculty Award.

References

Elliott-Graves, A., 2016. The problem of prediction in invasion biology. *Biology and Philosophy*, *31(3)*, 373–393.

Gorman, M., editor, 2010. *Trading Zones and Interactional Expertise*. MIT Press, Cambridge.

Griesemer, J., 2018. Mapping theoretical and evidential landscapes in ecological science. In Jeschke, J. and Heger, T., editors, *Invasion Biology: Hypotheses and Evidence*, pages 23–29, CABI, Boston.

Hardwig, J., 1985. Epistemic dependence. *Journal of Philosophy*, *82(7)*, 335–349.

Heger, T. and Jeschke, J., 2018. The hierarchy-of-hypotheses approach updated—a toolbox for structuring and analysing theory, research and evidence. In Jeschke, J. and Heger, T., editors, *Invasion Biology: Hypotheses and Evidence*, pages 38–45, CABI, Boston.

Hlobil, U., 2018. Choosing your nonmonotonic logic: A shopper's guide. In Arazim, P. and Lávička, T., editors, *The Logica Yearbook 2017*, pages 109–123, College Publications, London.

Horty, J., 2012. *Reasons as Defaults*. Oxford University Press, Oxford.

Jeschke, J. and Heger, T., editors, 2018. *Invasion Biology: Hypotheses and Evidence*. CABI.

Jeschke, J., Lokatis, S., Bartram, I., and Tockner, K., 2019. Knowledge in the dark: Scientific challenges and ways forward. *Facets*, *4*, 1–19.

Levins, R., 1966. The strategy of model building in population biology. *American Scientist*, *54(4)*, 421–431.

Lewis, M., 1988. *Cane Toads*. Film Australia, Sydney.

Meadows, D. H., Meadows, D. L., Randers, J., and Behrens, W. W., 1974. *The Limits to Growth*. Universe/Potomac Associates, Falls Church, VA.

Millgram, E., 2008. Specificationism. In Adler, J. and Rips, L., editors, *Reasoning: Studies of Human Inference and its Foundations*, Cambridge University Press, Cambridge.

Millgram, E., 2015. *The Great Endarkenment*. Oxford University Press, Oxford.

Morar, N., Toadvine, T., and Bohannan, B., 2015. Biodiversity at twenty-five years: Revolution or red herring? *Ethics, Policy and Environment*, *18(1)*, 16–29.

Moulton, M., Cropper, W., and Avery, M., 2011. A reassessment of the role of propagule pressure in influencing fates of passerine introductions to New Zealand. *Biodiversity and Conservation*, *20*, 607–623.

Mullaly, K. and Allen, T., 2015. *Land of OR*. Faceted Press, Park City.

Nguyen, C. T., 2018. Expertise and the fragmentation of intellectual autonomy. *Philosophical Inquiries*, *6(2)*, 107–124.

Ravetz, J., 1979. *Scientific Knowledge and Its Social Problems*. Oxford University Press, New York.

Reutlinger, A., Schurz, G., and Hüttemann, A., 2017. Ceteris paribus laws. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University.

Santana, C., 2014. Save the planet: Eliminate biodiversity. *Biology and Philosophy*, *29*, 761–780.

Santana, C., 2018. Biodiversity is a chimera, and chimeras aren't real. *Biology and Philosophy*, *33*.

Schupbach, J., 2018. Robustness analysis as explanatory reasoning. *British Journal of Philosophy of Science*, *69*, 275–300.

Suter, M., 2000. *Business Class*. Diogenes, Zurich.

Tabery, J., 2014. *Beyond* Versus*: The Struggle to Understand the* Interaction *of Nature and Nurture*. MIT Press, Cambridge.

Weisberg, M., 2006. Robustness analysis. *Philosophy of Science*, *73*, 730–742.

Elijah Millgram is E. E. Ericksen Distinguished Professor of Philosophy at the University of Utah and author of *The Great Endarkenment* and *Hard Truths*.