

Bach on Default Reasoning and Externalist Epistemic Justification

**Jeff Buechner
Department of Philosophy
Rutgers University-Newark and
The Saul Kripke Center
CUNY, The Graduate Center**

In an ingenious defense of reliabilism, Kent Bach argues in his paper “A Rationale for Reliabilism,”¹ that default reasoning is externalist and that it is necessary for providing epistemic justification of a wide range of everyday beliefs, such as perceptual beliefs. However, his description of what default reasoning consists in leaves open a number of important questions, and his discussion of internalism in epistemology suggests, but does not explicitly pose, a difficult problem. My purpose in this paper is (i) to show that formal epistemology has an important job in both providing formal explications of notions that appear in informal descriptions of default reasoning and in making precise what are the limitations of default reasoning, and (ii) to describe a new problem for internalism, and how default reasoning solves that problem. If Bach is right that default reasoning is externalist, then we are left with an interesting, if not bizarre, view of internalism—it requires externalism to be coherent. Finally, I will suggest a formal model analogous to the Tarski hierarchy in terms of which the problem for internalism is resolved, and conjecture a relation between this model and a solution to one version of the frame problem—a version which closely parallels Austin’s famous problem for epistemic justification.

Bach on default reasoning and the epistemic justification of everyday beliefs

A serious problem for internalist accounts of epistemic justification is that many beliefs which fail to have an internalist justification are, *prima facie*, cases of knowledge. Bonjour writes that “any non-externalist account of empirical knowledge that has any plausibility will impose standards for justification which very many beliefs that seem commonsensically to be cases of knowledge fail to meet in any full and explicit fashion. And thus on such a view, such beliefs will not strictly speaking be instances of adequate justification and of knowledge. But it does not follow that externalism must be correct. This would follow only with the addition of the premise that the judgments of common sense in this area are sacrosanct, that any departure from them is enough to demonstrate that a theory of knowledge is inadequate.”² Bach claims that almost all of our everyday beliefs—those that we acquire on the basis of perception of the particular environment we inhabit—are instances of knowledge, and, thus, that they are epistemically justified. But how are they justified? Bach thinks that the justification of those everyday beliefs must be sensitive to the epistemic and cognitive limitations on the reasoning which is about the situations which form the content of those beliefs.³ That is, we form an extraordinary number of everyday beliefs very quickly. If, for each belief, we had to provide an explicit chain of reasoning in order to epistemically justify that belief, we would not be able to justify many beliefs. The question, then, is how these beliefs get justified without an explicit chain of reasoning. Of course, the issue could be avoided by appealing to some notion of basic belief and formulating a causal story about

how perceptual mechanisms that result in the acquisition of basic beliefs are reliable. But it is a solution that is not satisfactory, since many of our everyday beliefs are conditional in the sense that they depend on other beliefs, whether basic or not. Moreover, some, such as Alvin Goldman,⁴ have argued that it is nonsense to think that there are regulative doxastic principles (because there would be a fundamental circularity in formulating such principles on the basis of the kinds of beliefs to which they are sensitive) and consequently, that one should look, instead, for doxastic habits which would epistemically justify everyday beliefs (whether basic or conditional).

It is an ingenious insight of Bach to see that a doxastic habit that would epistemically justify both basic and conditional everyday beliefs consists in the application of default reasoning in inferring those beliefs. How does default reasoning solve the limitation problem that explicit reasoning encounters? In default reasoning, steps are taken by default. That is, one jumps to a conclusion unless there is a reason not to jump to that conclusion. The important feature of default reasoning, as Bach conceives of it, is that it is not merely jumping to a conclusion, period. If default reasoning were the kind of reasoning, then it would hardly qualify as a means of conferring epistemic justification upon the beliefs at which one arrives using it. For, jumping to a conclusion without considering any kind of evidence for that conclusion is no better than arriving at a belief on the basis of flipping a coin. Without considering any evidence for the belief, we have no more reason to believe that it is true than we have reason to believe that it is false.⁵ Rather, default reasoning, in order to confer epistemic justification upon the beliefs that are inferred using it, must contain provisos about what happens when conditions are not ordinary.

Bach formulates what he dubs the take-for-granted principle (TFGP) in the following way:

“(TFGP) Its appearing to one that p justifies directly inferring that p provided that (a) it does not occur to one that the situation might be out of the ordinary, and (b) if the situation were out of the ordinary, it probably would occur to one that the situation might be out of the ordinary.”⁶

He claims—though without any proof—that beliefs that are arrived at on the basis of TFGP are reliable; that is, such beliefs are epistemically justified. How is it that the use of TFGP is not an internalist form of justification? The idea is that there are inferential steps in using TFGP that are not explicitly taken, viz., the steps that are described by conditions (a) and (b) in the formulation of TFGP. It is in not having a thought that things might be out of the ordinary and in the truth of the counterfactual, that if things were out of the ordinary, then it would occur to one that they are out of the ordinary, that one avoids having to make explicit inferences. Rather, one simply infers p, period. In inferring that p is the case, the conditions (a) and (b) are not explicitly followed. That is exactly where the default conception of reasoning is externalist. Conditions (a) and (b) do not come explicitly to mind in making the inference

that p is the case. These conditions are not explicitly expressed or represented in one's reasoning; rather they are implicit in the reasoning. But what does that mean?

Implicit assumptions in reasoning

Bach attempts to provide a definition of when something is implicit in one's reasoning by distinguishing between realizing an inference pattern and instantiating an inference pattern. "A piece of reasoning realizes an abstract pattern of inference if it contains psychologically real elements corresponding to all the steps of that pattern. It merely instantiates that pattern if there is some step that is not explicitly included but merely implicitly assumed."⁷ The problem, though, is that this definition does not really tell us what it is to be implicitly assumed, and thus what it is to be implicit in one's reasoning, since the notion of being implicit is used to define the notion of instantiating an inference pattern. How important is it that there is a notion of what it is to be implicitly assumed in one's reasoning that can be cashed out in terms of something—a computable mechanism, perhaps? For one, if it is left undefined as a primitive notion, then in arguing that default reasoning is reliable, the reliability of any given piece of default reasoning will be hostage to what it is for a piece of reasoning to implicitly assume something.

Suppose there are two individuals, A and B, who engage in default reasoning. A implicitly assumes that he can detect circumstances that are out of the ordinary, as does B. However, it is much harder for A to actually detect those circumstances, than it is for B. This has to do with how circumstances are represented in both A's and B's minds. A employs a representation that is cumbersome—determining that present circumstances are not normal is an NP-complete problem. However, B's representation is much more efficient—determining that circumstances are out of the ordinary is a problem in P, and, in fact, in one of the lowest levels of P. So although both A and B implicitly assume that they can detect when something is out of the ordinary, it is much easier for B to do it than it is for A. There will be circumstances, indeed, in which A does not have the resources to do it, while B does. In which case, the reliability of default reasoning will differ for A and for B. This might not be considered an epistemological problem, since the norms of default reasoning should not, one might argue, be sensitive to the resources of the agent. However, recall the motivation that Bach provides for default reasoning—"a reasonable theory of justified belief (hence of knowledge) must take into account real-life limitations on our everyday reasoning."⁸ It is default reasoning that precisely takes such limitations into account—we would not be able to infer quickly and in great quantities everyday perceptual beliefs unless we did engage in default reasoning. We simply jump to a conclusion that p , unless the thought that something is something is out of the ordinary occurs to us. So our episode of

default reasoning implicitly assumes that we can detect when something is out of the ordinary. But if we do not have an adequate notion of what it is for an inference pattern to implicitly assume something (such as a belief, a skill, and so on), then we cannot properly assess the reliability of default reasoning. If we cannot properly assess the reliability of default reasoning, then we do not have any argument for why default reasoning provides a rationale for reliabilism (or any other forms of externalist justificatory accounts).

Detection mechanisms and when they should be activated

Conceding that there is an adequate notion of what it is for a piece of reasoning to implicitly assume something, there are additional problems to which Bach's definition of TFGP succumbs. He says that "a belief resulting from such a process [TFGP] is justified to the extent that the process not only leads to true beliefs, but also guards against forming false beliefs, by means of precautionary subroutines that are generally activated when and only when they need to be. For it only to that extent that following TFG[P] can lead to justified beliefs."⁹ The problem is that it is not clear what he means by the phrase "when and only when they need to be." First, is he taking 'when and only when' to mean logical equivalence? If so, and if it means that the precautionary subroutine is activated when and only when one would, if it were not activated, infer a false belief, then it is much too strong. It would make the use of default reasoning infallible. That might be quite nice, but it is hardly realistic. What sort of detection mechanism must one have in place in order to be infallible in inferring one's perceptual beliefs by means of default reasoning? On the other hand, if the notion is loosened up a bit, how much loosening is permitted? That is, how often can one fail to engage a precautionary subroutine (and thus infer a false belief) and still be reliable in inferring true perceptual beliefs? This is a problem that will not go away. Perhaps Bach should use a cut-off that one employs in inductive reasoning: if the precautionary subroutine is engaged in N% of those cases (and $N \geq M$ —the cut-off) where one would otherwise infer a false belief, then it is reliable.

Internalist evaluations of reasoning

Bach contends that a problem for internalism is that there are aspects of reasoning that are generally not practically evaluable, since internalist reasoners who attempted to make such evaluations would find that they would be able to make only a few inferences. On the other hand, though a case could be made that reliability would increase were certain evaluations to be made that, the fact that making them would be impractical in terms of one's resources is a reason to not make them. But Bach claims that an important advantage of default reasoning over internalist reasoning is that the requirement of making such evaluations can be imposed differently for the externalist than it is for the

internalist. For externalist default reasoning, it can be imposed counterfactually: “to impose the requirement merely that the believer be able to perform such an evaluation would be plausible only if coupled with the requirement that the believer know when actually to perform the evaluation. However, the latter is a reliabilist requirement, not an internalist one.”¹⁰ That is, not only must episodes of default reasoning require that one be able to perform the evaluations of determining when something is out of the ordinary, but, additionally, the default reasoner must also know when something is out of the ordinary. However, this is too strong. For if one knows when something is out of the ordinary, then the sensitivity to circumstances being out of the ordinary must be epistemically justified—indeed, reliable. But how did this reliability arise? That is left a mystery. It also appears to generate a regress of reasons, in the following way. If the reliability of default reasoning depends upon the reliability of being able to determine when circumstances are out of the ordinary, then what does the reliability of determining when circumstances are out of the ordinary depend upon? It appears that detection of such circumstances is not merely a matter of perception, so that a causal story about how one is appropriately and causally connected to the local environment would not work. But, then, what will work? Is an appeal to default reasoning needed at this stage? If so, then the problem of the reliability of the detection mechanism for out of the ordinary circumstances will arise again, at a meta-level, and a regress of reasons is launched. On the other hand, if the appeal is not made to default reasoning, but to some form of reason that would justify the reliability of the detection mechanism in the ground-level instance of default reasoning, then the question of what makes that a reliable (justifiable) reasons arises, and the regress of reasons begins.

A dilemma for Bach

But there is a price for not imposing the requirement, as Bach explicitly notes: “... even if it were plausible to require merely that the believer be able to perform a thoroughgoing evaluation, that would divorce what justifies a belief from the process actually leading to it.”¹¹ That is, if the believer does not know when to perform the evaluation, but is only able to do it, the process that leads to the belief and in terms of which we say the belief is justified (externally) would be divorced from whatever it is that justifies the belief. However, Bach is firmly caught in a dilemma: if the knowledge requirement is in place, then there is either a regress of reasons or infallibility in detecting what is out of the ordinary. On the other hand, if it is not in place, then the notions of what justifies a belief and the process leading up to the justification of the belief are separated. This separation would doom externalist accounts of epistemic justification, since the externalist wishes to make sense of the notion of a belief that is justified even in cases in which the believer does nothing explicitly to justify the belief, and it is the

notion of a belief being justified in terms of the process leading up to that belief which allows externalism to succeed. But if it turns out that the price of success is that a requirement of knowledge is imposed on the detection mechanisms that are integral to default reasoning, it is not clear how externalist accounts of epistemic justification in terms of default reasoning can survive. Bach's project, which looked quite promising in resolving some of the tensions between internalist and externalist conceptions of epistemic justification is a non-starter.

Irrelevant alternatives

There is another problem to which Bach's conception of TFGP succumbs, which is closely related to the problem described above of what it is to implicitly assume a proposition in one's reasoning. Bach gives us another feature of this notion: "We implicitly assume a proposition whenever we reason in a way that is sensitive to it: drawing inferences consistent with it and not drawing ones inconsistent with it."¹² But under this feature of what it is to implicitly assume a proposition in one's reasoning, it is easy to show that one can implicitly assume almost anything, so long as it is not inconsistent with one's reasoning. This creates an enormous difficulty for Bach, since, as we have noted, a central motivation for introducing default reasoning as a form of externalist epistemic justification is that it provides an account of the kind of real-life reasoning that is needed to infer everyday beliefs about one's local environment. Here is the difficulty. Consider two agents, A and B, each of whom uses default reasoning to infer everyday perceptual beliefs. One way of cashing in what happens when one has a reason to think that things are not as they seem is to think of relevant alternatives to the hypothesis that they are as they seem.

Suppose the following counterfactual is true of A: if A had to (because things are not as they seem to be), A would think of the relevant alternatives RA_1 and RA_2 . Suppose the following counterfactual is true of B: if B had to (because things are not as they seem to be), B would think of RA_1 , RA_2 , IRA_1 , and IRA_2 (where IRA_1 and IRA_2 are irrelevant alternatives that are consistent with everything that B currently believes). The problem is this: both A and B implicitly assume the propositions expressed by RA_i and IRA_i . In each case, there is no inconsistency—so under Bach's definition of what it is to implicitly assume a proposition in one's reasoning, both A and B satisfy the definition. But there is an important difference between A and B—A implicitly assumes only what is necessary for the purposes of epistemic justification via default reasoning, while B implicitly assumes what is necessary and also what is unnecessary for the purposes of epistemic justification via default reasoning. We could easily extend the number of irrelevant alternatives that B implicitly assumes to an arbitrary number bounded from above by an ordinal of countable infinity. It is clear that B implicitly assumes too many

propositions. If circumstances were out of the ordinary, and the occurrence of a thought about those relevant alternatives were necessary for B to successfully engage in default reasoning, B would not be able to complete the episode of default reasoning in a reasonable amount of time. Thus, B would be no better off than an internalist reasoner who had to mentally represent all of the explicit steps in some chain of reasoning in virtue of which he infers perceptual everyday beliefs about his local environment. Either the notion of implicitly assuming a proposition in one's reasoning needs to be emended in order to rule out irrelevant alternatives being genuine instances of implicit beliefs, or else there has to be some set of constraints in place that will rule out irrelevant alternatives in descriptions of the conditions under which one does not jump to a conclusion in episodes of default reasoning.

There are certain kinds of frame problems in which the problem of irrelevance arises, notably those in which reasoning—concerning what action to perform—that examines irrelevant conditions will require too much time. How does a reasoning system determine what is relevant and what is not relevant without having to examine each instance and decide of it, that it is relevant or that it is irrelevant? The claim is the human beings are able to solve this problem, but how it is solved is an open question. Certain kinds of algorithms will not solve the problem, for they make the wasteful computations that examine every instance. The point is that Bach's TFGP achieves reliability only if there is in place some means by which irrelevant alternatives are not considered when situations are out of the ordinary and the precautionary subroutines are called into play. As we will argue below, it will be fruitful to examine formal work on solving the frame problem. Formal solutions to the frame problem may also be useful in providing a solution to the problem of irrelevant alternatives in TFGP.

If Bach does not rule out irrelevant alternatives in some way, then he loses the rationale for reliabilism that default reasoning provides, since the default reasoner will be no better off than the internalist reasoner, and perhaps will be worse off, if there are sufficiently many irrelevant alternatives that the default reasoner will make even fewer everyday perceptual inferences than the internalist reasoner. Of course, it might well happen that A and B (described in the preceding paragraph) will always jump to conclusions—there will never be occasions in which things are out of the ordinary. But although the default reasoner does better than the internalist reasoner because default reasoning takes into consideration that explicit steps in reasoning consume resources, and there is not enough time for explicit internalist reasoning in inferring everyday perceptual beliefs, the counterfactuals that are true of the default reasoner had better respect resource limitations, as well as establish the reliability of their inferences. If those counterfactuals do not respect resource limitations, then any episodes of default reasoning in which they figure are not a rationale for reliabilism, since they will not distinguish the

default reasoned from the internalist reasoner. Thus, it would be pointless to argue that there is a probability less than some cut-off value that both A and B will not encounter situations when things are out of the ordinary. Even if such out of the ordinary situations are not encountered, A will, but B will not, satisfy the requirements of externalist epistemic justification.

What does Bach mean by “out of the ordinary?”

The final problem for Bach is to specify what out of the ordinary situations consist in—that is, how it is that a default reasoner can recognize that things are out of the ordinary? This is obviously related to the problem of knowing when things are out of the ordinary, though here the emphasis is on what procedures or mechanisms are employed to recognize when things are out of the ordinary. Once they are recognized to be out of the ordinary, the default reasoner has a reason to think that he should not jump to a conclusion sanctioned by TFGP. It would appear that things can be out of the ordinary in both ways that matter for believing *p* and in ways that do not matter for believing *p*. The question is: which of these ways should a default reasoner be sensitive to? Once again, there is a resource limitations problem that threatens conditions (a) and (b) of TFGP. Here is how that happens: suppose that A never fails to recognize when things are out of the ordinary and that A is always justified in believing that things are out of the ordinary. Thus A knows when things are out of the ordinary. However, most of the situations that are out of the ordinary have nothing to do with the justification of *p*, where *p* is an everyday perceptual belief that is arrived at on the basis of a default inference using TFGP. In which case, A will do no better than an internalist reasoner, since A will get bogged down in actually considering that things are out of the ordinary and will not make as many everyday perceptual inferences as will a default reasoner who is less sensitive to what is out of the ordinary, but who always manages to recognize those out of the ordinary situations that are relevant to the justification of her everyday perceptual beliefs. Of course, Bach can build into TFGP that only those out of the ordinary situations that are relevant to the justification of *p* are to be considered. But now that creates an extraordinary difficulty—one that is reminiscent of attempts to provide necessary and conditions for the acquisition of mental contents.

The difficulty is that in determining which out of the ordinary situations are relevant to the justification of *p*, an agent must know what evidence counts in justifying *p* and what evidence does not count in justifying *p*. But in order to know what evidence does and does not count in justifying *p*, the agent must have necessary and sufficient conditions for what *p* means. That is, whatever information *p* expresses, the agent must know what it is, since, otherwise, she would not be able to determine what is and what is not relevant evidence for justifying *p*. There is nothing untoward about this—it is a simple

reflection of the obvious fact about reasoning that evidence for a claim must provide information that will eliminate counterexamples to that claim. The more evidence there is, the fewer counterexamples there are to the reasoning for that claim. In John Pollock's terminology, the agent must be sensitive to undercutting defeaters to what she infers by a default inference.¹² But to distinguish between genuine counterexamples (genuine undercutting defeaters) and irrelevant ones, the agent must know what the claim means. Otherwise, the agent will not know when there is a genuine undercutting defeater, since such a defeater is one in which the evidence holds, but the claim is false. Without having necessary and sufficient conditions for what a claim means (and thus what information it expresses), it is impossible to determine with any kind of reliability when there is a genuine undercutting defeater and when there is an irrelevant undercutting defeater—that is, something that is not an undercutting defeater, even though it is erroneously taken to be one.

There is an additional complication in determining when situations are not ordinary, and it is that there are qualitative differences between situations which are abnormal, but within normal bounds, and situations which are abnormal, but within abnormal bounds, which closely parallels the difference between situations which are normally abnormal and situations which are abnormally abnormal. This kind of difference between abnormal situations arises frequently in real life situations, such as medical diagnosis. What difference will it make to a default reasoner who is unable to distinguish between kinds of abnormal situations? One criticism is that it makes no difference, so long as the default reasoner is able to detect that the situation is abnormal. This, however, is false. It is well-known from work on non-monotonic reasoning that the Yale Shooter's Problem shows that there are important differences.¹³ Consider the following scenario. At a certain time, there is a loaded gun. Later, the gun is aimed at a person and fired. Subsequent to the firing of the gun, the person who is hit by the bullet dies. The situation is formalized in terms of John McCarthy's ontological notions of fluent, actions, effect axioms and observation axioms. Suppose that a reasoning system is designed to minimize abnormalities, as is the case in circumscriptive reasoning systems. The problem is that there are normal abnormalities and abnormal abnormalities. A normal abnormality is that—in the context of the Yale Shooter's Problem—someone who is alive becomes dead. An abnormal abnormality is that a loaded gun suddenly becomes unloaded. The problem the Yale Shooters Problem signals is that unless we are able to distinguish between normal and abnormal abnormalities, simply minimizing abnormalities may be fruitless, if we minimize the wrong abnormalities. Different proposals were advanced for how to solve the Yale Shooters Problem, the most promising of which took into account the temporal order of the abnormalities and the kinds of causal histories that would account for that temporal order. Here, too,

examining formal work in non-monotonic reasoning—such as various forms of circumscription—may be fruitful in formulating a more realistic TFGP.

The problems I have enumerated for Bach's TFGP show that in the absence of providing a formal account of TFGP, conditions (a) and (b) are nothing more than a mystery mechanism. It is easy to claim that TFGP will result in epistemically justified beliefs—that it is a reliable inference rule—but that claim is merely a promissory note that needs to be cashed in before any of Bach's interesting and substantial points about externalist accounts of epistemic justification can be taken seriously. There has been, since 1980, much work on formalizing default and inductive reasoning. It would not be possible to examine all of this work, with an eye toward providing an explication of Bach's TFGP. However, Reiter's formalization of default reasoning provides an important place to begin that task, since Reiter explicitly engages with the question of what it means for a situation to be out of the ordinary.

Reiter's formalization of default reasoning

In his paper "A Logic for Default Reasoning,"¹⁴ Ray Reiter introduced into the literature an important formalization of default reasoning. It is different in motivation from Bach's TFGP, since it is not concerned with making fast real life inferences, as in making everyday perceptual inferences. But it is concerned with reliability—that is, where one has less than complete information about a situation, how can reliability in one's inferences about that situation be established? This is an old problem, that of making reliable inductive inferences. The difference is that rather than modeling an inductive logic on deductive logic, the idea is build the logic for default reasoning in such a way that it is defeasible—that is, in such a way that inferences that are made can be taken back. The core idea is that defeasible inferences in a logic are non-monotonic. In classical logic, inferences are monotonic—the addition of information which is added to the system does not alter the status of what has already been proved within the system. That is, given $A \vdash B$, it logically follows that $A, C \vdash B$, for any C . Inferences in classical logic are local—they do not depend upon what else happens in the system. On the other hand, defeasible reasoning is not local. What else that happens in the system can determine the status of a formula—whether it is provable or not.

In Reiter's system of default reasoning, he provides an interpretation of what it means for a situation to be out of the ordinary—when there is information to the contrary. His interpretation is that there the situation is normal—there is no information to the contrary—when it is consistent to jump to the default conclusion. Suppose the default conclusion is that Tweety, a bird, flies. The default inference rule that sanctions this default conclusion has the form: Premise $BIRD(x) : MFLY(x)$; Conclusion: $FLY(x)$. The term 'M' is that it is consistent to assume that x flies. If it is not consistent to assume that x flies,

then the default conclusion that x flies is not inferred. If at some later stage in a chain of default reasoning, it becomes known that it is not consistent to assume that x flies, then the default conclusion that x flies is retracted. The problem, then, is to specify the set of conditions that are necessary to establish the assumption that it is consistent to assume that x flies. Obviously, consistency with everything that the reasoner believes is infeasible, since the set of beliefs the agent harbors might be quite large. Indeed, determining the consistency of 'x flies' with everything that the reasoner believes is an NP-complete problem, since one would have to examine all subsets of the set of N beliefs to determine the consistency of 'x flies' with the entire set of beliefs. 'x flies' might be consistent with belief A and consistent with belief B , but it might be inconsistent with the conjunction of beliefs A and B . There is an additional problem: if the reasoner reasons about sentences of arithmetic, then the entire system of reasoning—the default system and the first-order deductive system that contains a theory of arithmetic, cannot prove the consistency of 'x flies' with the rest of the sentences that are provable in the combined system of reasoning.

No doubt, for these reasons, Reiter must restrict the sentences that are within the scope of the M predicate. In the case of inferring by default that Tweety flies, there are explicit exceptions to the claim that x flies which are: Tweety is a penguin or Tweety is an ostrich. $M \text{ Fly}(x)$ is consistent with $BIRD(x)$ and $FLY(x)$ provided that neither $PENGUIN(x)$ nor $OSTRICH(x)$ are believed by the default reasoner. Suppose that $OSTRICH(x)$ is believed by the default reasoner. There is also the following rule in the system: $(x) OSTRICH(x) \rightarrow \sim FLY(x)$. If the default reasoner believes that $OSTRICH(\text{Tweety})$, then by universal instantiation, $OSTRICH(\text{Tweety}) \rightarrow \sim FLY(\text{Tweety})$ and modus ponens, $\sim FLY(\text{Tweety})$. But $\sim FLY(\text{Tweety})$ is inconsistent with $FLY(\text{Tweety})$. Thus $MFLY(\text{Tweety})$ is false, and the default conclusion that Tweety flies is not inferred, or retracted (if it has already been inferred).

The problem in interpreting the notion of what it is for things to be out of the ordinary as consistency with some set of beliefs is that all of the beliefs that are relevant to establish consistency must be explicitly included in the default logic. Consider how many exceptions to the claim that Tweety flies there might be: Tweety is an ostrich, a penguin, a stuffed bird, a roasted bird, a toy bird, a sick bird, a dead bird, and so on. However, even if all the exceptions are listed, there are still complications: one is that Tweety might well be a toy bird, but a toy bird that is equipped with wings and a motor that enables it to fly, though for only very short periods. In this case, there are exceptions to the exceptions. Should these exceptions be taken into account as well? One proposal would be that one defaults to the conclusion that Tweety flies unless it is believed that Tweety is a toy bird, and then, unless it is believed that Tweety is a toy bird equipped with wings and a motor. That is, there will be default inferences

about the exceptions to the default inference. This will create a hierarchy of default inferences, which should not be confused with interacting defaults that produce conflicts in what is inferred by default. Here the hierarchy is organized in terms of a root default inference and leaves in a tree that represent the default inferences that are made about the exceptions that are necessary for plausibly interpreting the M predicate.

The question, then, is whether Reiter's formalization of a system of default reasoning solves any of the problems that arise for Bach's TFGP. The advantage of Reiter's formalization is that it gives a precise meaning to "situation that is out of the ordinary." In fact, formalizing default reasoning is important for just this reason: we need to know whether a proposed condition for a default inference to not be made is plausible or feasible or possible. In interpreting "situation that is out of the ordinary" as "it is consistent to assume that ... ," we learn what is feasible, possible, and plausible. However, the consistency interpretation is one of many different interpretations. Moreover, it clearly does not answer all of the problems that arise for Bach's TFGP. In particular, the issue about irrelevant situations is not answered. Different formalizations of non-monotonic reasoning provide different instruments with which to explore problems that arise in the description of default reasoning. In that way, they can be useful in examining epistemological questions, such as what reliability in externalist theories of epistemic justification consist in. They can also provide upper bounds on what can be done in a system of default reasoning.

A new problem for internalist theories of epistemic justification

One way of defining an internalist conception of epistemic justification is that it "requires a person have a 'cognitive grasp' of whatever makes his belief justified. Being justified depends on how rational and 'epistemically responsible' (whatever these mean precisely) he is in coming to hold the belief."¹⁵ In arguing the default reasoning is externalist, Bach assumes that reasoning that results in epistemic justification in an internalist sense must make all of its steps explicit. Thus, default reasoning is not internalist in this sense since some steps are not taken. That is, the description of default reasoning in, for example, TFGP, cannot be explicitly followed. There are respects in which TFGP could be evaluated, but in most instances is not. These respects concern the conditions describing what happens when situations are out of the ordinary.

Bach suggests, but does not describe, a problem for internalist conceptions of epistemic justification that might be unsolvable on internalist grounds. Consider what he says about a valid argument in some system of logic. "When our reasoning to a conclusion is sufficiently complex, we do not survey the entire argument for validity. We go more or less step by step, and as we proceed, we

assume that if each step follows from what precedes [it], nothing has gone wrong. ... Even if our lines of reasoning were always perspicuous, so that we could view them as a whole, there would still be points at which we do not actually check for validity but simply 'go along' with the reasoning at that point. We just 'see' that the next step follows. In any case, to lead to justified belief reasoning does not have to be evaluated in every evaluable respect."¹⁶ What he is suggesting, though not explicitly describing, is a problem for internalism: if internalist reasoning must evaluate all of the respects in which reasoning which results in internalist epistemically justified beliefs can be evaluated, how can it evaluate in every evaluable respect without succumbing to the Lewis Carroll problem for modus ponens? Recall that the problem Carroll posed for logical inferences, such as modus ponens, is how one follows the rule of modus ponens without using some other rule which will justify the steps that one follows in following modus ponens. In particular, how does one avoid using modus ponens to justify that one is actually following the rule of modus ponens?

There have been various responses to Carroll's problem, all of which endorse some form of externalist solution. For instance, Ned Block argued that the procedure for following modus ponens is simply hard-wired into the brain (or hard-wired into a computer chip).¹⁷ The idea is that it is not possible to evaluate a chain of reasoning in every evaluable respect without either assuming the use of such reasoning in one's evaluative assessment of that reasoning or else having the evaluations end in virtue of some hard-wired physical capacity (i.e., the employment of the rule is hard-wired). If there is no solution to Carroll's problem that is not externalist, then it, *prima facie*, is the case that internalist reasoning that epistemically justifies a proposition *p* necessarily requires externalism. Internalism is incapable, on its own terms, of ever providing epistemic justification of a proposition *p* via some chain of reasoning. If so, this is surprising. Even Bach does not take internalism and externalism to be connected intrinsically. His view is that there are two notions of justified belief, and one answers to internalist demands, while the other answers to externalist demands—this distinction is explicit in his notions of justified beliefs and justified believers.¹⁸

On the surface, the problem for internalism is remarkably like Austin's evidence problem: when do I have enough evidence to be justified in my belief that *p*? When is enough *enough*?¹⁹ The problem for internalism is: when I have made enough evaluations of my reasoning to epistemically justify the proposition that *p*? When do I stop making evaluations of my reasoning? However, the similarity is illusory, provided one takes a hard line about internalist epistemic justification. A hard line view about internalist epistemic justification is that until every aspect in which the reasoning for a proposition *p* can be evaluated is, in fact, taken, there is no epistemic justification. However, if one takes a more plausible

line, that epistemic justification occurs in the absence of total evaluation of the reasoning, then we have Austin's evidence problem. There is also a remarkable similarity between Austin's evidence problem and at least one version of the frame problem: how does a reasoner know when to stop looking at ways in which actions are qualified and come to a decision about the consequences of an action? Below I will propose a solution to the problem for internalism. It is a solution that will only work for those who take a hard line about internalism, though. It will not solve Austin's evidence problem.

If default reasoning is externalist, then a solution to the problem for internalism is to employ default reasoning in evaluating a chain of reasoning in every evaluable respect. Here is my proposal. Suppose that one is confronted with a valid chain of reasoning for some conclusion. Certainly, it is imperative to examine the chain of reasoning and to ascertain that it is an instance of some valid argument form. But once one has done that, there is no further need to evaluate what one has done in recognizing that the chain of reasoning is itself a substitution instance of some valid argument form. Rather, one assumes that the situation of recognizing that the chain of reasoning is a substitution instance of a valid argument form is normal and jumps to the conclusion that it is a substitution instance of a valid argument form. If the situation were out of the ordinary, then one would have a reason to think that the exhibition of that recognitional capacity is epistemically inappropriate. One might even use Bach's TFGP in carrying out this task. Since TFGP is externalist, one is clearly using externalist resources in order to provide an internalist epistemic justification of some proposition p .

If the default reasoning that is used in this context is formalized, say in Reiter's system of default reasoning, then it becomes possible to describe precisely any kinds of limitations in the internalist justification of p . Reiter's system of default reasoning has been shown to be non-cumulative—that is, there are certain non-monotonic rules of inference that are violated in his system. In particular, the non-monotonic rules that are violated are inclusion, supraclassicality, cautious monotony, and cut. What is the importance of knowing that, for instance, non-monotonic cut is violated? It is this: It has been shown that proofs that do not allow a cut rule can grow exponentially. (The exact form of their growth has been characterized by Alessandra Carbone, in a series of important papers.²⁰) Recall that default reasoning—in TFGP—is viewed by Bach as a way to make everyday perceptual inferences quickly and reliably. But if cut is violated, then chains of default reasoning might have more steps than either reasoning carried out in some other kind of non-monotonic logic or purely internalist reasoning.

Consider cautious monotony: if y non-monotonically follows from x and z non-monotonically follows from x , then infer that z non-monotonically follows from both x and y . This is violated in Reiter's system of default logic. Suppose that an internalist uses Reiter's default logic to infer that a chain of

reasoning is valid reasoning. It might be the case that the chain of reasoning is complex, in which case there is a need to show that several different parts of the chain of reasoning are substitution instances of some or other valid argument form. If we can avail ourselves of the non-monotonic rule of cautious monotony, then we can consolidate some of the work that has been done, so that we do not have to duplicate certain results. But this duplication will need to be done where cautious monotony is violated. Once again, saving inferential steps in a chain of inferences is a benefit of reasoning with TFGP, but what formalization shows is that this benefit fails to materialize in at least one case. The various forms of circumscription are cumulative, and so the non-monotonic inference rules noted above are non-monotonically valid in these forms. This suggests that where one's demands are efficiency in the use of computational resources, then some form of circumscription is better than default logic. To reiterate: the various formalizations of non-monotonic reasoning provide a set of tools that epistemologists can sue to examine default reasoning, and to examine the various ways of interpreting the conditions in, for instance, TFGP. Is a condition that states that the reasoner should not jump to p when there are relevant alternatives to p equivalent—and in what way(s) equivalent—to a condition that states that the reasoner should not jump to p when the situation appears to be out of the ordinary? Formal epistemology provides the tools to make such comparisons precise and meaningful.

It might be objected that using, say, Reiter's default logic to provide an internalist epistemic justification of some proposition p via a chain of reasoning that is not formalized in a logic violates the view of reasoning promoted by Gilbert Harman—that reasoning is not the same as making logical inferences.²¹ Isn't the use of Reiter's default logic not giving reasons that the internalist needs, but rather making a non-monotonic logical inference, which is not giving a reason? However, the use of, say, default logic in providing the internalist with the reasons she needs to provide an epistemic justification of some proposition p does not violate Harman's division of reasoning and logical inferences, since it is in the context of providing a reason for the internalist that the logical inference is made. That a logical inference is made does not mean that it cannot also serve as a reason—provided that the context is clear that a reason is what is being given. What is the reason for believing that the chain of inferences that epistemically justifies a proposition p on internalist grounds? The reason is that the set of all evaluations of the chain of reasoning that could be made has been closed off in virtue of making a default inference in Reiter's default logic.

The default inference is that the conclusion of the chain of deductive reasoning is validly inferred from its premises provided that there is no reason to think that the situation is out of the ordinary. But why can't the internalist question, where there is a reason to think that the situation is out

of the ordinary, that very reason? If that is permitted, then there will be a regress of reasons launched on the default inference, which is supposed to break the regress of reasons that arises on the internalist justification of the chain of deductive reasoning. In effect, in asking whether the reason that the situation is out of the ordinary is a good reason, one is asking whether the reason is reasonable. That is, might there not be a defeater of the reason—in particular, might it not be possible to find some situation in which all of the data which serve as evidence that the situation is out of the ordinary is preserved, yet the conclusion that the situation is out of the ordinary is denied? The problem is that, although such a defeater might appear incoherent (provided that the data is correctly assessed), it cannot be ruled out of bounds on internalist principles. Indeed, that the reasoner might not understand the data, or might not be able to correctly recognize that the data is in fact credible and warrants the conclusion are problems for a non-ideal internalist reasoner. Where a reasoner is susceptible to error, arising from various contingent sources, such as a faulty memory or too much caffeine or a headache, they may mistakenly conclude that the reason for thinking the situation is out of the ordinary is not a good reason, and is subject to defeaters. Of course, the externalist would respond that the data, in and of itself, does justify the conclusion. Now the internalist could go along with the externalist on this, provided that they distinguish between contingent sources of error that are abnormalities in normal functioning of a human being and contingent sources of error that are normal limitations in the functioning of a human being. Thus, having too much caffeine may result in becoming mentally ‘hyper,’ and finding defeaters where there are none. On the other hand, if the description of the situation is too complex, it may not be possible to keep all of the data in short-term memory. The latter is a contingent limitation in the functioning of a human being, while the former is a contingent abnormality in the normal functioning of a human being.

The point is that if making a distinction between normal and abnormal sources of error is the only way in which the internalist can avoid what is a regress of reasons, then internalism is hostage to that distinction. Drawing it in a principled and non-question begging way might be difficult, since there are surely cases in which it is not clear whether something is a matter of normal human functioning or an abnormality in normal human functioning. In any case, making an appeal to a conception of normal human functioning in order to save internalism from collapsing into a doctrine in which it is impossible to epistemically justify any claim does not really prevent the collapse, if in defining normal human functioning the aim is to prevent the collapse. Here, however, is where a default conception of reasoning can play a role. Since the problem the internalist creates is that the reason for thinking that a situation is out of the ordinary, which defeats the default inference, may itself be defeated, why not

take a meta-level default inference on the object-level default inference? Of course, the internalist will argue that the very same problem arises at the meta-level. In which case, one ascends to a higher level. It is easy to index a iterated chain of default inferences of default inferences. The formal question that arises is when does the chain close off? It appears likely that, just as a truth predicate in a Tarskian theory of truth will close off at the ordinal ω_1^{CK} , so too will the chain of default inferences of default inferences. Clearly, no finitary human being can follow the chain out to ω_1^{CK} , but that hardly matters. Rather, the formal result would tell us that the chain of default inferences over default inferences closes off at that stage in the inductive construction. Knowing that, we are justified in believing that the original chain of deductive reasoning—which the internalist cannot evaluate in all evaluative respects—is justified. That is, the transfinite iteration of default inferences of default inferences shows us that nothing out of the ordinary will happen. Of course, the internalist could find problems for the transfinite iteration—recall the discussion of sources of error above. But the point is that it is the externalist transfinite construction of a sequence of default inferences of default inferences that justifies the internalist reasoning that the chain of deductive reasoning is justified. Given any finite stage in the transfinite construction of the sequence of default inferences of default inferences, the internalist can raise an objection that it is not justified. But the externalist only has to ascend to a next higher level to answer the objection. Thus, at any finite stage, the internalist objection that the construction is not justified can be defused by the externalist, who always has recourse to ascend to a higher stage in the inductive construction. In this way, it is established that default reasoning, under the supposition that it is externalist, is necessary for the completion of internalist justification of some proposition p .

Fodor's fridgeons

In an important and somewhat neglected paper, “Frames, Fridgeons, ,”²² Jerry Fodor argues that the kinds of problems addressed by non-monotonic logics are all the problem of induction in disguise. Whether Fodor is right about that is not my concern, though. Rather, it is his introduction of fridgeons that I would like to consider, in the context of a problem for Bach’s TFGP that I described earlier, that of irrelevant information. Bach is mindful of this problem: “Our inferences must take relevant information into account without getting bogged down in irrelevancies.”²³ But he nowhere proposes how to solve that problem. Once again, it is a problem for an externalist account of epistemic justification, since the conditions (a) and (b) in TFGP are needed for the reliability of default reasoning and they easily succumb to irrelevancies unless additional constraints are imposed on them to rule out the irrelevancies.

Fodor attributes to Drew McDermott a solution to one kind of frame problem: when an action occurs, most things do not change.²⁴ How do you formally specify that most things do not change

without simply listing all of the things that do not change (when an action occurs)? McDermott's intuitive notion is that one does not let sleeping dogs awaken. This is characterized formally by an axiom of persistence. Fodor's complaint is that McDermott's solution will not work when certain predicates are defined and true of objects in the world. Fodor defines the predicate of being a fridgeon as: x is a fridgeon at time t_0 if and only if Fodor's refrigerator is on at t_0 . Suppose that one adopts the axiom of persistence. Then only those things that are explicitly proved to change as a result of an action change; everything else stays the same. That, however, will not be the case when fridgeon predicates are introduced into a world description. As soon as Fodor turns on his refrigerator (an action), every physical particle in the world changes its state, from being a non-fridgeon to being a fridgeon. Moreover, it is true of any physical particle, after Fodor turns on his refrigerator, that they have the property of being fridgeons. Thus, the predicate 'fridgeon' is unlike the predicate 'grue.' The latter creates a well-known problem for inductive reasoning—the same evidence base confirms competing inductive generalizations. But in Fodor's fridgeon construction, there is no competition with other inductive generalizations. What Fodor has done is create a procedure for generating irrelevant truths about the world. Obviously, no reasonable person would think about the world in terms of such predicates; but that does not say anything about how to rule them out of world descriptions and one's reasoning about the world. It is also important to note, though Fodor does not make this point, that one acquires knowledge about the world in justifying beliefs about the fridgeon status of physical objects. Thus, there is an interesting kind of converse skeptical problem: how do we avoid knowing too much about the world? How do we acquire knowledge that is appropriate for living in the world and do not acquire knowledge that is not appropriate for living in the world?

Fodor's construction can obviously be generalized in various ways, all of which result in a nightmare of irrelevant truths and knowledge claims. For instance, one could define fridgeon-like predicates that would require much effort in determining whether they are instantiated in physical objects. (x is a szhlepton if and only if the fly on my windowsill moves .0000005 meters to its left.) It is patent that fridgeons are irrelevant predicates and that truths about fridgeons are irrelevant truths. Imagine that a world creator is creating our world and wants to make sure that there are enough truths about the physical world for us to discover. The world creator reaches a stage in the creation process where there are countably many truths about the physical world and wonders whether she should add more. She thinks: "Is enough *enough*? Shouldn't I add a few more truths? How do I know that enough is *enough*?" Fodor's fridgeon problem can be viewed, then, as another instance of Austin's "How do I know when enough is *enough*?" problem.

Although this problem infects conditions (a) and (b) in TFGP, could formal solutions to the characterization of default reasoning illuminate the problem? Recall that the internalist conception of justified reasoning engendered a problem—that of knowing when to stop evaluating evaluable aspects of one’s reasoning. When is enough *enough* in evaluating one’s reasoning? The suggestion that iterated transfinite sequences of default reasoning about default reasoning will solve the internalists’ problem (though the solution is externalist, given that default reasoning is externalist). The suggestion for solving Fodor’s fridgeon problem is that TFGP is used, given that there is a conception of what are the normal predicates used to describe the world. At that point, one jumps to the conclusion that, unless the situation appears to be out of the ordinary, that is all of the predicates that are needed to describe the world. The default conclusion would not be available when new predicates are introduced into the world description, such as ‘vulcanize.’ The science of vulcanization and the resulting technology would be sufficient to trigger conditions (a) and (b) in TFGP. Similarly, in acquiring evidence for one’s claims about the world, default reasoning can be employed to defuse the problem of when enough evidence is *enough*. The default conclusion is that there is enough evidence—this would require that there are well established guidelines for sufficient evidence for claims of various kinds. The guidelines for different claims might differ accordingly. Conditions (a) and (b) are triggered whenever there is reason to think that there is not sufficient evidence. Finally, should there be internalist worries about the very reasoning itself (at the meta-level), these can be assuaged by creating iterative transfinite sequences of default reasoning about default reasoning, an inductive construction that closes off at the ordinal ω_1^{CK} .

Conclusion

Bach has argued that default reasoning provides an interesting and important way of understanding externalist conceptions of epistemic justification. However, his description of default reasoning succumbs to several difficult problems, all of which endanger his epistemological project. I have pointed out several of these problems. My view is that they can be fruitfully examined by looking at different formalizations of non-monotonic reasoning. In this paper, I looked at only one such formalization—Reiter’s default logic. There are various notions that are used in filling out principles like TFGP, such as “out of the ordinary,” abnormal, alternative possibility, and so on. Formalizations of non-monotonic reasoning can provide useful formal characterizations of these notions and their interrelationships. They can also provide upper bounds on what reasoning in these systems can accomplish (such as the existence of extensions for default logic). This paper is only a first step in doing the work that will (one hopes), show what is needed for externalist conceptions of epistemic

justification to be viable and to go beyond something more than mere talk of reliable mystery mechanisms.

FOOTNOTES

1. Kent Bach, "A Rationale for Reliabilism," *Monist*, 1985: 246-263.
2. Lawrence Bonjour, "Externalist Theories of Empirical Knowledge," *Midwest Studies in Philosophy*, 5 (1980): 97-111, p. 97.
3. That is a view that might be rejected by many epistemologists, since it comes close to bringing psychologism into epistemology. Those who reject psychologism argue that cognitive limitations should not play a role in the description of epistemic norms. This is an important issue, though I won't take a stand on it here, and will agree with Bach that one way of looking at everyday perceptual beliefs is that they are instances of knowledge, and that given our cognitive limitations, an account of how we acquire that knowledge is necessary.
4. Alvin Goldman, "The Internalist Conception of Justification," *Midwest Studies in Philosophy*, 5 (1980) : 27-51, p. 37-38.
5. If there is no more evidence for than against a belief, than one should be indifferent whether to believe p or believe not-p, provided that the evidence does not sanction a probability of p being true of greater than or less than .5, and similarly for the evidence for not-p.
6. Bach, op. cit., p. 259.
7. Bach, op. cit., p. 257.
8. Bach, op. cit., p. 255.
9. Bach, op. cit., p. 258-259.
10. Bach, op. cit., p. 256.
11. Bach, op. cit., p. 256.
12. John Pollock, *How To Build A Person*, MIT Press, Cambridge, 1989, p. 126.
13. Murray Shanahan, *Solving the Frame Problem*, MIT Press, Cambridge, 1997, pp. 60ff.
14. R. Reiter, "A Logic for Default Reasoning," *Artificial Intelligence*, 13 (1980) : 81-132.
15. Bach, op. cit., p. 247.
16. Bach, op. cit., p. 257.
17. Ned Block, "The Mind as Software of the Brain", *Thinking An Invitation to Cognitive Science* Volume Three (Second edition), MIT Press, Cambridge, 377-425.
18. Bach, op. cit., pp. 251-253.
19. J. L. Austin, *Sense and Sensibilia*, Oxford University Press, 1962, p. 125ff.

20. Alessandra Carbone, "Duplication of Directed Graphs and Exponential Blow Up of Proofs," *Annals of Pure and Applied Logic*, 100 (1999 : 1-76).
21. Gilbert Harman, *Change in View*, MIT Press, 1986, Chapter 1.
22. Jerry Fodor, "Modules, Frames, Fridgeons, Sleeping Dogs, and the Music of the Spheres," *Modularity in Knowledge Representation and Natural Language Understanding*, Jay Garfield (ed.), MIT Press, Cambridge, 1987, 26-36.
23. Bach, op. cit., p. 258.
24. Drew McDermott, "We've Been Framed: Or, Why AI Is Innocent of the Frame Problem," *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*, Ablex, 1986.