

Deliberation, Responsibility, and Excusing Mistakes of Law

Alexander A. Guerrero
University of Pennsylvania

I. Introduction

In *Excusing Mistakes of Law*, Gideon Yaffe sets out to “vindicate” the claim “that mistakes of law never excuse” by “identifying the truth that is groped for but not grasped by those who assert that ignorance of law is no excuse.”¹ Yaffe does not offer a defense of the claim that mistakes of law *never* excuse. That claim, Yaffe argues, is false. Yaffe’s article is, rather, an effort to assess what plausible thought might be behind the idea that mistakes of law often *should not* excuse. (Yaffe is interested in more than just the descriptive claim that in Anglo-American legal jurisdictions mistakes of law routinely do not, in fact, excuse.) More particularly, Yaffe is interested in what plausible *normative* justification there might be for this asymmetric pattern:

ASYMMETRY: False beliefs about non-legal facts often excuse, but false beliefs about the law rarely excuse.

Yaffe offers a complex argument in support of ASYMMETRY. This paper is organized around my reconstruction of Yaffe’s argument. I will argue that Yaffe’s argument does not succeed, but that his argument may provide a template for an argument that could succeed.

II. Deliberation as the Root of Responsibility

Here are the first two premises of Yaffe’s argument:

(P1) The capacity to deliberate is a necessary feature for the distinctive kind of moral censure that is only appropriate for persons.

(P2) Features of an agent’s deliberation “serve as the grounds of her responsibility”; specifically, “corruption in deliberation . . . is the root of responsibility for wrongdoing.”²

Yaffe acknowledges that these claims are controversial, but he states that “they are plausible enough to warrant assuming” as true.³

It is worth noting that even if one accepts (P1), one might not accept (P2). That is, one might hold that moral censure is only appropriate for those entities that deliberate without holding that the root of responsibility for wrongdoing must be *corruption* in deliberation. It may be that moral censure only makes sense when aimed at creatures that have control over what they do (in some sense of control), and the ability to deliberate is a necessary condition of having control of this sort. But one might still maintain that an agent is responsible for the action simply because the action was the agent’s action, that it was something that they intended to do. On such a view, it could be that the agent deliberated well, but that (in Yaffe’s terms) the inputs were faulty—the agent began with a perverse set of moral beliefs or

¹ Gideon Yaffe, “Excusing Mistakes of Law,” *Philosophers’ Imprint*, 9 (2009), 1.

² Yaffe, 10.

³ Yaffe, 10.

values, or non-moral factual beliefs, and this led the agent to behave badly, and in a way for which (it is natural to say) he is responsible.

Now, in some cases, it may be that bad deliberation is what led the agent to have these misguided moral beliefs, values, or non-moral factual beliefs, and so although it is the inputs that led the agent astray, the *root* of responsibility is the bad deliberation that led to the agent accepting these beliefs or values in the first place. But this need not be the case. It could be that the agent deliberated well all along the way, but his evidence was such that it led him to these false or perverse beliefs and values, despite deliberating well. (Perhaps he was raised in a sheltered or warped or prejudiced environment.) Some might wish to excuse such individuals, perhaps because they accept (P2). Others might have more of a ‘strict liability’ approach, maintaining that all that must be true for an agent to be morally responsible for an action is that the action was an action of hers—her beliefs and deliberations led her to intend to do the action, and she successfully acted on that intention. On such a view, if one has false beliefs, and these beliefs lead one—via perfectly fine deliberation—to act badly, one is still morally responsible for acting badly, even if it is not one’s fault that one had the false beliefs in the first place.

These are old and familiar issues. For any action, X, performed by some person, Q, there are a host of factors that led Q to do X. Some of these have to do with what Q believes and values, some of them have to do with how Q deliberates. And there are likely other subconscious or non-rational factors that led Q to do X. For all of these factors, there is a causal story about how they came to obtain—how one came to have these beliefs and values, how one learned to deliberate using one’s beliefs and values, how one came to be in these particular decisionmaking circumstances, etc. The old and familiar question is *how much of all this must one have had control over, or be responsible for, or have caused, in order for one to be morally responsible for doing X?* Galen Strawson, in offering what he calls the Basic Argument, pushes the line that in order to be morally responsible for what you do, you must have been ultimately responsible for being the way you are—you must have been ultimately responsible for all or almost all of these factors.⁴ According to Strawson, because it is clear that one cannot be ultimately responsible for these factors, one cannot be morally responsible for what one does.

Strawson’s argument is a skeptical one, leading to the view that no one is morally responsible for anything. To resist the argument, one must find some stopping place, some factor which can ground moral responsibility, despite it being the case that the agent is not ultimately responsible for everything that led up to that factor obtaining. In endorsing (P2), Yaffe implicitly takes *how an agent deliberated* to be a factor of this sort. But he offers no argument for why this factor is the appropriate one. The importance of this becomes starkly apparent with the next premise of his argument:

(P3) If one’s deliberations to perform some action X were *uncorrupted*, then one’s decision to do X does not deserve punishment; if one’s deliberations to perform some action X were *corrupted*, then one’s decision to do X does deserve (might deserve, often deserves) punishment.

This premise is the heart of his normative defense of ASYMMETRY. Yaffe assumes that *how an agent deliberated* prior to deciding to do X is relevant to moral⁵ desert, and can serve as a basis for moral

⁴ Galen Strawson, “The Impossibility of Moral Responsibility,” *Philosophical Studies* 75 (1994), 5.

⁵ Later in the paper, Yaffe offers a distinction between legal desert and moral desert, and says of an individual who has “made a mistake about an arbitrary legal rule” that “she deserves [punishment] legally and does not deserve it morally,” but that what is wrong about punishing her is “not that she does not deserve the punishment she receives.” Yaffe, 20. Yaffe says that “[t]he degree to which people legally deserve punishment for choosing to act in a way unfavored by the balance of legal reasons is the degree to which such a choice indicates a commitment to faulty principles for the recognition and weighting of legal reasons.” *Id.* But the challenge for those who would defend Asymmetry is to say why it is *morally* appropriate to punish such a person—what appears to be wrong with punishing such a person is precisely that she does not morally deserve the punishment she would receive. Yaffe’s argument cannot possibly succeed if (P3) is read as being about only *legal* desert of punishment. The first part of

responsibility. Equally importantly, Yaffe assumes that an agent's *beliefs* that entered into the decision to do X are *not* relevant to moral desert, cannot serve as a basis for moral responsibility, as long as the agent's deliberations pertaining to doing X—using those beliefs—were uncorrupted.⁶

These two assumptions require defense. In particular, it is not obvious, given Strawson-style concerns about moral responsibility, that *how an agent deliberated* is more morally central for desert than *what an agent believes*, particularly when those beliefs can include beliefs about morality, what is valuable, and non-moral facts related to those questions. For one, it seems that how we deliberate—and in particular how good we are at deliberation—can be equally out of our control (on some natural understanding of control) as what we believe. And both seem relevant to ordinary assessment of moral character. It is certainly not obvious that we care more, morally, about how agents *deliberate* than about what they *believe*—given that both of these play a central role in what we really care about: what agents *do* (outside just their heads).

III. Principled Deliberation

It is worth considering Yaffe's view of what deliberation is, and how it can be corrupted, to see if this will help make sense of this pair of assumptions. That view is expressed in the next two premises:

(P4) Whether one's deliberations were corrupted depends on whether one followed correct *principles* regarding (a) what reasons to consider (rules for extracting reasons from facts) and (b) what weight to give those reasons (rules for weighting the reasons one has extracted).

(P5) Whether a principle is *correct* depends on whether the principle yields correct results—that is, a principle for extracting reasons from facts is correct just in case the facts in question do indeed provide the reasons that the principle specifies.

Yaffe is explicit that by “principles” he does not mean “formal rules like the means-end principle,” but rather “general rules for extracting reasons from facts and general rules for weighting the reasons one has extracted.”⁷ Here is an example of a principle of the first sort: “If an act is likely to cause a person pain, then that is a reason against performing that act.”⁸ Here is an example of a principle of the second sort: “If one act is more likely to cause a person pain than another, then weigh the reason against performing the first more heavily than the reason against performing the second.” He notes that in the legal context, “we will be concerned only with principles of this kind confined to legal reasons . . . principles for extracting *legal* reasons from facts, and the principles for weighting *legal* reasons the agent being evaluated is committed to.”⁹ Yaffe states that “deliberation” on his account is “the norm-governed psychological process through which a person reaches a decision and so often forms an intention to act in a particular way.”¹⁰

I confess to finding this notion of principle, and the notion of principled deliberation it suggests, somewhat opaque. I have three main difficulties with Yaffe's account of deliberation. First, the notion of

(P3) is false on such a reading, and the argument as a whole would become just about what is legally appropriate, rather than about what is morally appropriate with respect to excuse and punishment for mistakes.

⁶ Yaffe does mention the issue of whether a person who is responsible for having a false belief can be denied an excuse on the grounds of having that false belief. He declines to engage with that issue, noting that “[w]e need only find a principle of excuse that meets the four criteria and which could be made true by adding the requirement that the defendant not be culpable for having made the mistake.” Yaffe, 6. I think that this issue is more central than Yaffe's response acknowledges, a point that I will make in Section V.

⁷ Yaffe, 11.

⁸ *Id.*

⁹ *Id.*

¹⁰ Yaffe, 10.

deliberation via principles—in this sense of principles—is idiosyncratic. It is natural to think of deliberation as involving the formation and revision of beliefs in light of evidence and by using inductive and deductive principles of reasoning, and the judging and assigning values or weights to various options in lights of one’s beliefs, values, and preferences. Some of this is the province of epistemology, some of rational choice theory, some of ethics. How to deliberate well is a complex issue. One would expect that the relevant normative principles about such matters would be fairly general. But that is not the picture that Yaffe has in mind. He discusses the following example (throughout, my presentation of examples are slightly modified for ease of reference):

MARITAL RAPE: A man, Williams, has non-consensual sex with his wife, falsely believing that there is a marital exemption from rape in the jurisdiction in which he lives.

Factual Beliefs	Legal Beliefs
Williams F1: “My wife is not consenting to have sex in this particular instance.” [True]	Williams L1: “One is legally permitted to have non-consensual sex with one’s wife in this legal jurisdiction.” [False]
	Williams L2: “I am within this legal jurisdiction.” [True]
	Williams L3: “I am legally permitted to have non-consensual sex with my wife.” [False]

Yaffe says of this man that his false beliefs about the law indicate “a commitment on his part to the following principle for extracting reasons from facts: his wife’s lack of consent provides no legal reason against having sex with her.”¹¹ If this is a “principle” of deliberation, it is a principle of a very specific kind. It does not seem particularly well-described as a principle at all. Even if we re-describe the principle as “whether one’s wife consents to sex or not does not affect what legal reasons there are” or something of the sort, it is still a very fine-grained kind of principle. Consider another case that Yaffe discusses at length:

CHILD CARE: Battersby entered into a contract to house and take care of a couple’s children five days a week, with the children returned to their parents on the weekends. Battersby did not have a license to provide foster care. There was a statute that required a person to be licensed if caring for children for money for more than thirty consecutive days. In Battersby’s case, a court ruled that, for the purposes of the statute, Friday and Monday are consecutive days—in effect, the court ruled that thirty consecutive *weekdays* of care required a license.

Factual Beliefs	Legal Beliefs
Battersby F1: “I am not caring for children for more than 5 consecutive calendar days in a row.” [True]	Battersby L1: “Friday and Monday are not, for legal purposes, consecutive days.” [False]
	Battersby L2: “My contractual arrangement with the couple does not require me to obtain a license.” [False]

Yaffe says that Battersby’s false beliefs about the law “show her to be committed to the following principle for extracting reasons from facts: the fact that one is providing care for more than thirty consecutive weekdays does not provide one with a legal reason to get a license.”¹² Again, this is nothing like a general principle of deliberation, or a general principle for extracting reasons from facts. Yaffe

¹¹ Yaffe, 13.

¹² *Id.*

appears to envision an incredible multitude of “principles of deliberation.” This is odd, but this might be no more than a terminological point about the use of the phrase “principles of deliberation” were it not for the next problems.

The second problem is a straightforward concern about equivocation: the notion of corrupt deliberation that emerges in (P4) and (P5) does not seem like the same notion of corrupt deliberation that might make (P2) and (P3) true. Recall that (P2) maintains that “corruption in deliberation is the root of responsibility for wrongdoing” and (P3) states that “if one’s deliberations to perform some action X were corrupted, then one’s decision to do X does deserve (might deserve, often deserves) punishment.” Whatever one might have thought of these premises with a more intuitive understanding of deliberation, they seem much harder to accept with the specialized notion of deliberation that Yaffe employs. As noted above, one might have a view of deliberation on which deliberation involves behavior governed by epistemic, prudential, and moral norms, so that one might fail to deliberate well in a number of different ways. For example, before deciding whether or not to demolish a building, one might deliberate weighing economic costs and benefits and legal zoning regulations without considering whether there are any living creatures in the building that might be harmed. This might be a moral failure of deliberation—one has failed to investigate a question relevant to the moral permissibility of one’s action prior to acting. Or one might have considered that question, but given the answer insufficient weight in deciding what to do. Failures of deliberation in this respect might be the right kind of thing to make (P2) and (P3) seem plausible. It is much harder to see that “corrupted deliberation” in the sense identified above, so that the man who does not think that his wife’s consent is relevant to the legal reasons he has is *deliberating* badly, could ground (P2) and (P3).

This is particularly true if we limit our attention to deliberation via principles for extracting *legal* reasons from facts. Perhaps we think that certain generally deliberatively defective individuals deserve moral censure for their deliberative failings. Maybe they never take other people’s interests or experiences of pain into account when deciding what to do. But it is not plausible that we feel this way about faulty deliberation using these kinds of principles for extracting *legal* reasons from factual circumstances. Consider again the man who rapes his wife, who engaged in faulty deliberation because he accepts this principle for extracting legal reasons from facts: his wife’s lack of consent provides no legal reason against having sex with her. It is implausible that any part of the reason the man deserves punishment, or at least meets the precondition of moral desert, is that his deliberation was corrupted because he acts on this principle of extracting *legal* reasons from facts. He deserves punishment because he forced another person to have sex with him, or because he took what legal reasons he believed he had to exhaust the moral reasons present—not because he deliberated badly about what legal reasons he had.¹³

The third concern about Yaffe’s notion of deliberation and deliberative principles is that, upon closer examination, these *principles of deliberation* just seem to be straightforward *beliefs* about reasons or *beliefs* about value. Yaffe says that his principles are “general rules for extracting reasons from facts and general rules for weighting the reasons one has extracted.”¹⁴ But, as noted above, the principles he discusses are not general at all. Maybe that isn’t so important. What does seem important is that a distinction can be drawn between “the inputs to the deliberative machine” and “the machine itself.”¹⁵ As Yaffe puts it:

¹³ The somewhat obscure distinction Yaffe makes between “what rules the law establishes” and “what legal reasons the law provides” is of no help. Yaffe, 13-14. Yaffe suggests that what is relevant for assessing whether a principle is accurate or not is whether it tracks the legal reasons, not the actual legal rule (as determined, say, by a legal authority like a court). How exactly to make out this distinction is unclear, but Yaffe says that “The fact that the statute *in fact*, as the court rules, directs her [to get a license] does not imply, on this view, that she has a legal reason to get one: the legal reasons on this view are supplied by the most natural interpretation of the statute and not by the facts about what rule the law establishes, when the two diverge.” Yaffe, 14. But neither of these two possible measuring sticks for principles help to make bad deliberation, corrupted deliberation, deliberation on the basis of bad principles, look any more relevant to moral desert.

¹⁴ Yaffe, 11.

¹⁵ Yaffe, 10.

The guiding assumption in what follows is that successful excuses sometimes operate by indicating that the agent's deliberation was uncorrupted in itself. When we offer an excuse of this sort we admit that the agent's conduct was objectionable; we would prefer a world that didn't contain such conduct. But we imply that while something went wrong with the inputs to the deliberative machine, the machine itself worked as it ought. We are claiming that uncorrupted deliberation yielded a corrupted outcome; but since it is corruption in deliberation that is the root of responsibility for wrongdoing, an excuse is warranted in such a case.¹⁶

But, although the distinction is vital for his account, it is not clear that this distinction between the machine and the inputs to the machine can be drawn, or that Yaffe draws it in an intelligible way. Yaffe contends that all of the following are deliberative principles that are *part of the machine* rather than beliefs that are *inputs to the machine*:

“his wife's lack of consent provides no legal reason against having sex with her”¹⁷

“the fact that one is providing care for more than thirty consecutive weekdays does not provide one with a legal reason to get a license”¹⁸

“a bar order gives one no legal reason not to enter a military base if one is entering in order to participate in a peace demonstration”¹⁹

“the fact that an unconstitutional statute says that a person must do something gives him no legal reason to do it”²⁰

Why are these principles part of the deliberative machine, rather than inputs to it? This concern gives additional bite to the equivocation worry: even if we can draw the distinction somehow, it is hard to see how it could ground anything as strong as (P3). In particular, there does not seem to be any normatively significant difference between our ability to control, or our responsibility for, which *principles* we are committed to and our ability to control, or our responsibility for, what we *believe*, if control or responsibility are supposed to be why bad deliberation is damning but deliberating with bad inputs is not. Nor does deliberation via principles of this sort seem more deeply connected to our character, or who we are, than what we believe. This puts additional pressure on Yaffe's commitment to deliberation as the place where the responsibility buck stops. What is special about deliberation, if this is what we mean by deliberation?

IV. Bad Beliefs v. Bad Deliberation

Here are the next premises of Yaffe's argument, in which he continues to place significant weight on the distinction between mistakes that result from bad deliberation and mistakes that result from bad inputs:

¹⁶ *Id.*

¹⁷ Yaffe, 13.

¹⁸ *Id.*

¹⁹ Yaffe, 15.

²⁰ Yaffe, 16.

(P6) Some false beliefs can excuse because they imply that, although the agent performed the wrong action, her deliberation was proper and uncorrupted; the wrong action was performed because the agent had a false belief, not because the agent deliberated improperly.

(P7) If an agent has a false belief that *p*, and this false belief *indicates* that in her deliberations pertaining to doing some action *X*, the agent was committed to correct principles for extracting reasons from facts and correct principles for weighting reasons, then the agent is excused for doing *X*.

(P8) False beliefs that indicate that one is committed to incorrect principles do not exculpate; they inculcate.

In the previous section, I raised a number of concerns about the distinction that Yaffe draws between deliberative inputs (beliefs) and principles of deliberation. An even more immediate problem can be raised for this approach. How, in a given case, are we to tell which of the following two situations an agent is in when an agent ends up acting wrongly?

EXCUSABLE: (false belief + fine deliberation) → wrong action
 INEXCUSABLE: (false belief + bad deliberation) → wrong action

I have framed this as a question of how we know whether to attribute principles of deliberation to agents in a particular case. Leaving aside the epistemic concern, there is a metaphysical concern: what makes it the case that an agent is in one of these situations or the other?

Consider the example with which Yaffe begins the paper:

WIRE: With his landlord’s permission, Smith installed some speaker wire behind a wall in an apartment he was renting. When he moved out, he disconnected the wire. He believed that the speaker wire belong to him, although the law is that things attached *to* a wall belonged to the tenant; things installed behind walls belong to the landlord.

Factual Beliefs	Legal Beliefs
Smith F1: “The wire is behind the wall.” [True]	Smith L1: “One is legally prohibited from taking things that do not belong to one when one moves out of a rented property.” [True]
Smith F2: “I acquired and installed the speaker wire.” [True]	Smith L2: “One is legally permitted to take what one owns when one moves.” [True]
	Smith L3: “One has legal ownership over those things—even those things behind walls—which one acquires and installs.” [False]
	Smith L4: “The speaker wire belongs to me.” [False]
	Smith L5: “I am legally permitted to remove the wire from the wall and to take it with me.” [False]

Is Smith an EXCUSABLE or INEXCUSABLE case? Consider the following mistaken principle of deliberation: the fact that one is moving a wire from behind a wall provides no legal reason to obtain permission from the landlord. Is Smith committed to this principle, so that his is an INEXCUSABLE case? Or does he act wrongly only because he falsely believes Smith L3, that “[o]ne has legal ownership over those things—even those things behind walls—which one acquires and installs,” so that his is an EXCUSABLE case? If one believes Smith L3, does this belief *indicate* a commitment to the mistaken principle? Yaffe says that it was right to excuse Smith because of his mistaken belief about the law. But why? It seems that his false legal belief indicates a commitment to a mistaken principle of deliberation in

exactly the same way that the marital rapist’s false legal belief indicates a commitment to a mistaken principle of deliberation.

Or consider the Cheek case that Yaffe discusses:

TAXES: Cheek attended seminars hosted by groups interested in dramatically reforming the American tax system. Hearing from these speakers, Cheek came to believe that wages are not income for tax purposes, and that tax laws are unconstitutional. As a result of these beliefs, Cheek failed to pay any income tax for years.

Factual Beliefs	Legal Beliefs
	Cheek L1: “Wages are not income for tax purposes.” [False]
	Cheek L2: “Federal income tax laws are unconstitutional.” [False]
	Cheek L3: “Unconstitutional laws are not legally binding.” [False]
	Cheek L4: “I don’t need to report wages for tax purposes.” [False]
	Cheek L5: “I don’t need to pay taxes.” [False]

Yaffe says that Cheek should not be excused because he was committed to the following mistaken principle: “the fact that an unconstitutional statute says that a person must do something gives him no legal reason to do it.”²¹ But how do we know whether Cheek is committed to this principle? Why not think that he just falsely believes Cheek L3, and that Cheek L3 is serving as a bad input to perfectly fine deliberative machinery? What makes it the case that Cheek was committed to this principle (or not)? This makes all the difference on Yaffe’s account, but it seems completely arbitrary (and normatively insignificant) whether we opt for one description of his case or the other.

Here is a more general way to make the point. Consider the following conjecture:

CONJECTURE: In every case in which an individual has a false belief, that false belief is an indication that the agent was committed to an incorrect principle of deliberation, and that her deliberation was to that extent corrupted.

Given the understanding of deliberation and principles of deliberation that Yaffe advances, it seems that CONJECTURE is true. We can always *construct* an incorrect principle of deliberation that generates the false belief, for any false belief. Given that agents do not *actually* reason using the principles that Yaffe attributes to them (consider the cases of Battersby or the marital rapist), it is hard to see what constraint there is on this sort of construction. This is a problem for Yaffe’s account, since the final steps in his argument for ASYMMETRY are these:

- (P9) False non-legal beliefs *rarely* indicate that one is committed to incorrect principles.
- (P10) False normative beliefs *often* indicate that one is committed to incorrect principles.
- (P11) In particular, false *legal* beliefs often indicate that one is committed to incorrect principles.
- (P12) False legal beliefs often, and false non-legal factual beliefs only rarely, implicate defects in our deliberative mechanisms that ground our desert of punishment.
- (P13) Therefore, it is normatively appropriate that ASYMMETRY is true; that false beliefs about non-legal facts often excuse, but false beliefs about the law rarely excuse.

²¹ Yaffe, 16.

But if CONJECTURE is true, then all of (P9)-(P12) are false.²² Most significantly, the apparent difference between false legal and non-legal factual beliefs expressed in (P12) disappears.

What might Yaffe say to resist CONJECTURE? The obvious response is to show that false non-legal beliefs do not provide any indication, or do not as regularly provide any indication, that the agent was committed to an incorrect principle of deliberation; no construction of such a principle is possible. It does not seem that such a response will work. It is certainly not obvious that we have reason to think that there is some systematic difference in what factual and legal false beliefs, respectively, reveal about our commitments to incorrect principles. Consider the following paradigmatic case of false non-legal belief.

COFFEE: Adam is bringing coffee to Brianne. Unbeknownst to him, the sugar dish has been emptied and filled with arsenic. He puts a spoonful of what he believes to be sugar, but what is in fact arsenic, into Brianne's coffee. She drinks the coffee and dies.

Adam has a false non-legal factual belief: that what is in the sugar dish is sugar. He needn't have any false legal views about the permissibility of putting arsenic in a person's coffee or anything of the sort. Is Adam's false non-legal belief indication that he is committed to any incorrect principles? What about this one: the fact that one is giving a person a substance to ingest does not always provide one with a reason to make sure that the substance is safe to consume. In general, principles of this sort will be able to be constructed because one can always suggest that the person is committed to a principle that licensed the particular level of epistemic inquiry that the person engaged in, even though that level of inquiry resulted in the person making a mistake.

Is Adam committed to this principle? Again, there is the question: how would we know? But it seems no odder to say that Adam is committed to this principle than it is to say that Battersby is committed to this one: "the fact that one is providing care for more than thirty consecutive weekdays does not provide one with a legal reason to get a license."

Is this principle correct? It seems that it is not. Recall that the standard of correctness for principles is demanding:

(P5) Whether a principle is *correct* depends on whether the principle yields correct results—that is, a principle for extracting reasons from facts is correct just in case the facts in question do indeed provide the reasons that the principle specifies.

In this case, the principle leads Adam astray, suggesting that he had *no* reason to make sure that the granular substance was safe to consume, whereas it seems that Adam does have a reason to make sure the substance is safe to consume, but that reason is (perhaps) outweighed by considerations of convenience. Of course, the question of whether he actually had such a reason is to wade into significant debates, including the internalism v. externalism debate, something that I don't want to do here. Whatever we think about this question, it seems that the most perspicuous criticism of Adam, if any criticism is appropriate (it would depend on unspecified details of the case), is not that he deliberated using an incorrect principle. If Adam is subject to criticism, it is not because he deliberated badly, but because he (in some sense) *should have known* that the substance was arsenic or *should have inquired* into whether the substance in the sugar dish was safe to consume—he is culpable for the false belief, or culpable for acting on that false belief, given what he had done to ascertain whether or not the substance he gave Brianne was arsenic. I will pursue this line of thought in the final section.

V. The Grain of Truth

²² (P10) and (P11) are false if "often" implies "often, but not always."

Early in the paper, Yaffe suggests that whether one is culpable for having the false belief—whether legal or factual—is not relevant to his immediate project of identifying a principle of excuse that satisfies his four criteria. That is because, as he puts it, “[w]e need only find a principle of excuse that meets the four criteria and which could be made true by adding the requirement that the defendant not be culpable for having made the mistake.”²³ The principle that Yaffe finds is the Uncorrupted Deliberation Principle, which is basically (P7):

(P7) If an agent has a false belief that *p*, and this false belief indicates that in her deliberations pertaining to doing some action X, the agent was committed to correct principles for extracting reasons from facts and correct principles for weighting reasons, then the agent is excused for doing X.

Yaffe’s suggestion is that we could modify this principle as follows:

(P7*) If an agent has a false belief that *p*, *and the agent is not culpable for believing p*, and this false belief indicates that in her deliberations pertaining to doing some action X, the agent was committed to correct principles for extracting reasons from facts and correct principles for weighting reasons, then the agent is excused for doing X.

Importantly, (P7*) has the effect of shrinking the set of cases of false belief under consideration. The question is whether (P9) and (P11) still seem true, as frequency claims, even when the set of cases has been shrunk in this way. So, suitably modified:

(P9*) False non-legal beliefs that one holds non-culpably *rarely* indicate that one is committed to incorrect principles.

(P11*) False legal beliefs that one holds non-culpably *often* indicate that one is committed to incorrect principles.

My intuition is that, whatever one might have thought of these quasi-statistical claims previously, it is very hard to defend them now. More precisely, it is hard to defend them with a notion of “incorrect principles” that grounds an assessment of morally blameworthy, bad deliberation. If (P9*) and (P11*) seem true, it is just because we think that the “non-culpable” proviso in (P9*) ensures that one is committed to using correct principles. If this is not true in (P11*), it is only because “incorrect” in that premise just means legally incorrect—not incorrect from the point of view of one’s epistemic or other deliberative obligations. But deliberation that is bad in that sense just isn’t morally bad, at least not when we have stipulated that the person holds the false legal beliefs non-culpably.

Yaffe suggests that the Uncorrupted Deliberation Principle “allows us to identify the grain of truth in the false slogan that ignorance of law never excuses.”²⁴ What I want to suggest is that there is a grain of truth in this vicinity, but it does not have to do with deliberation, or the distinction between mistakes made because of bad deliberation and mistakes made because of bad inputs.

On the reinterpretation of the account I am offering, what we would say instead of (P6) and (P7) is this:

(*) If *D* has a false belief that *p* and (a) *D*’s having that false belief explains why *D* performed some objectionable action X, (b) the fact that *D* believed that *p* is not itself grounds for finding *D* morally objectionable, and (c) the fact that *D* believed that *p* does

²³ Yaffe, 6.

²⁴ Yaffe, 12.

not indicate that *D* has failed to live up to morally-based epistemic obligations regarding the maintenance and formation of her beliefs, then *D* is excused.

Why should we prefer (*) to (P6) and (P7)? There is not space to defend (*) in detail here. One reason to prefer (*) is that it keeps us from having to draw Yaffe's difficult to sustain distinction between principles of deliberation and beliefs that serve as inputs to deliberation. A more significant reason is that (*) seems to both deliver and explain the right results with respect to a wide range of cases. It is intuitively plausible that Battersby and Smith should be excused, that Williams (the marital rapist) should not be, and that Cheek is somewhat more in the middle. Battersby and Smith should be off the hook because their false beliefs were reasonable (satisfying (b) and (c)), and those beliefs clearly explain why they behaved badly in the relevant sense (satisfying (a)).

Williams's false beliefs do not satisfy (a). He had false beliefs about the *law*, but those beliefs were irrelevant to the moral permissibility of his action, and do not explain why he had non-consensual sex with his wife. It is not as if, had the law permitted non-consensual sex, his action would have been morally permissible. Additionally, if he had the false belief *that if having non-consensual sex with his wife is legally permissible, then it is morally permissible*, this might well indicate that either (b) or (c) does not hold. In general, when (b) does not hold, it may be for the same reason that the fact that one is venal cannot be cited as an excuse for one's venal actions.

Cheek is somewhat more in the middle because although (a) and (b) are satisfied (if we assume that Cheek actually had the false beliefs), the way the case is described it is not clear whether (c) is satisfied. Given that there may be a moral duty to pay one's taxes, and given the moral significance of tax-paying, it is at least plausible Cheek may not have done enough to satisfy his moral obligations to investigate the question of whether he had a duty to pay taxes. And so it is unclear whether (c) is satisfied.

There seem to be two ways that (*) or some principle like it would vindicate ASYMMETRY. The first would be if we endorsed something like a kind of strict liability moral requirement to have correct beliefs about the law, so that having false beliefs about the law almost always meant that condition (b) was not satisfied. This does not seem plausible. Why should we have a strict liability moral requirement to have correct beliefs about the law, particularly given the complex and arcane nature of law in modern political systems? Note that this might be true with respect to certain core moral beliefs, so that merely having false beliefs about those topics would in effect constitute being morally objectionable.

The second way that (*) would vindicate ASYMMETRY is if the epistemic demands with respect to knowing the law (avoiding false beliefs about the law) were systematically higher than the epistemic demands with respect to knowing other factual matters—and higher not because of epistemic reasons, but because of moral ones. I argue for a related view elsewhere.²⁵ This would mean that (c) was less frequently satisfied with respect to false legal beliefs than with respect to false beliefs about non-legal matters. There might be a story that could be told in this vicinity, perhaps based on a Waldronian story about the importance of respect for legislation and democratically-created law.²⁶ Additionally, this approach would avoid holding individuals to a standard of perfection with respect to what they must learn and know about the law, while acknowledging something of the special significance to law that is implicit in Yaffe's (P11).

²⁵ See Alexander A. Guerrero, "Don't Know, Don't Kill: Moral Ignorance, Culpability, and Caution," *Philosophical Studies*, 136 (2007): 59-95.

²⁶ See Jeremy Waldron, *Law and Disagreement* (1999).