

TRUE ENOUGH*

Catherine Z. Elgin
Harvard University

Epistemology valorizes truth. Sometimes practical, or prudential, or political reasons convince us to accept a known falsehood, but most epistemologists deny that we can have cognitively good reasons to do so. Our overriding cognitive objective, they maintain, is the truth, preferably the whole truth, and definitely nothing but the truth (Goldman 1999, p. 5; Lehrer 1986, p. 6; Bonjour 1985, p. 9). If they are right, then at least insofar as our ends are cognitive, we should accept only what we consider true, take pains to insure that the claims we accept are in fact true, and promptly repudiate any previously accepted claims upon learning that they are false. I suggest, however, that the relation between truth and epistemic acceptability is both more tenuous and more circuitous than is standardly supposed. Sometimes, I contend, it is epistemically responsible to prescind from truth to achieve more global cognitive ends.

At first blush, this looks mad. To retain a falsehood merely because it has epistemologically attractive features seems the height of cognitive irresponsibility. Allegations of intellectual dishonesty, wishful thinking, false consciousness, or worse immediately leap to mind. But science routinely transgresses the boundary between truth and falsehood. It smoothes curves and ignores outliers. It develops and deploys simplified models that diverge, sometimes considerably, from the phenomena they purport to represent. Even the best scientific theories are not true. Not only are they plagued with anomalies and outstanding problems, but where they are successful, they rely on laws, models, idealizations and approximations that diverge from the truth. Truth-centered epistemology, or *veritism*, as Alvin Goldman calls it, easily accommodates anomalies and outstanding problems, since they are readily construed as defects. The problem comes with the laws, models, idealizations, and approximations which are acknowledged not to be true, but which are nonetheless critical to, indeed constitutive of, the understanding

that science delivers. Far from being defects, they figure ineliminably in the success of science. If truth is mandatory, much of our best science turns out to be epistemologically unacceptable and perhaps intellectually dishonest. Our predicament is this: We can retain the truth requirement and construe science either as cognitively defective or as non-cognitive, or we can reject, revise, or relax the truth requirement and remain cognitivists about, and fans of science.

I take it that science provides an understanding of the natural order. By this I do not mean merely that an ideal science *would* provide such an understanding or that in the end of inquiry science *will* provide one, but that much actual science has done so and continues to do so. I take it then that much actual science is cognitively reputable. So an adequate epistemology should explain what makes good science cognitively good. Too strict a commitment to truth stands in the way. Nor is science the only casualty. In other disciplines such as philosophy, and in everyday discourse, we often convey information and advance understanding by means of sentences that are not literally true. An adequate epistemology should account for this as well. A tenable theory is a tapestry of interconnected sentences that together constitute an understanding of a domain.¹ My thesis is that some sentences that figure ineliminably in tenable theories make no pretense of being true, but are not defective on that account. If I am right, theories and the understandings they embed have a more intricate symbolic structure than we standardly suppose. Nevertheless, I do not think that we should jettison concern for truth completely. The question is what role a truth commitment should play in a holism that recognizes a multiplicity of sometimes conflicting epistemological desiderata.

Consensus has it that epistemic acceptability requires something like justified and/or reliable, true belief. The justification, reliability, and belief requirements involve thresholds. '[O]ne may need to be confident enough and well enough justified and one's belief must perhaps derive from a reliable enough source, and be little enough liable to be false' (Sosa 2000, p. 2). But truth, unlike the other requirements, is supposed to be an absolute matter. Either the belief is true or it is not. I suggest, however, that the truth requirement on epistemic acceptability involves a threshold too. I am not saying that truth itself is a threshold concept. Perhaps such a construal of truth would facilitate treatments of vagueness, but that is not my concern. My point is rather that epistemic acceptability turns not on whether a sentence is true, but on whether it is true enough—that is, on whether it is close enough to the truth. 'True enough' obviously has a threshold.

I should begin by attempting to block some misunderstandings. I do not deny that (unqualified) truth is an intelligible concept or a realizable ideal. We readily understand instances of the (T) schema:

- ‘Snow is white’ is true \equiv snow is white
- ‘Power corrupts’ is true \equiv power corrupts
- ‘Neutrinos have mass’ is true \equiv neutrinos have mass

and so on. A disquotational theory of truth suffices to show that the criterion expressed in Convention (T) can be satisfied. One might, of course, want more from a theory of truth than satisfaction of Convention (T), but to make the case that the concept of truth is unobjectionable, a minimalist theory that evades the paradoxes suffices. Moreover, not only does it make sense to call a sentence true, we can often tell whether it is true. We are well aware not only that ‘Snow is white’ is true \equiv snow is white, but also that ‘Newly fallen snow is white’ is true. The intelligibility and realizability of truth, of course, show nothing about which sentences are true or which truths we can discover. Nevertheless, as far as I can see, nothing about the concept of truth discredits veritism. Since truth is an intelligible concept, epistemology can insist that only truths are epistemically acceptable. Since truth is a realizable objective, such a stance does not lead inexorably to skepticism. I do not deny that veritism is an available epistemic stance. But I think it is an unduly limiting one. It prevents epistemology from accounting for the full range of our cognitive achievements.

If epistemic acceptance is construed as belief, and epistemic acceptability as knowledge, the truth requirement seems reasonable. For cognizers like ourselves, there does not seem to be an epistemically significant gap between believing that p , and believing that p is true. Ordinarily, upon learning that our belief that p is false, we cease to believe that p . Moreover, we consider it cognitively obligatory to do so. One ought to believe only what is true. Perhaps a creature without a conception of truth can harbor beliefs. A cat, for example, might believe that there is a mouse in the wainscoting without believing that ‘There is a mouse in the wainscoting’ is true.² In that case, the connection between believing that p and believing that p is true is not exceptionless. But whatever we should say about cats, it does not seem feasible for any creature that has a conception of truth to believe that p without believing that p is true. If epistemic acceptance is a matter of belief, acceptance is closely linked to truth. Assertion is too. Although asserting that p is not the same as asserting that p is true, it seems plain that one ought not to assert that p if one is prepared to deny that p is true or to suspend judgment about whether p is true; nor ought one assert that p is true if one is prepared to deny that p or to suspend judgment about whether p . Assertion and belief, then, seem committed to truth. So does knowledge. Whether or not we take knowledge to be equivalent to justified or reliably generated true belief, once we discover the falsity of something we took ourselves to know, we withdraw the claim to knowledge. We say, ‘I thought I knew it, but I was wrong,’ not ‘I knew it, but I was wrong’.

Being skeptical about analyticity, I do not contend that a truth commitment is part of the meanings of ‘belief,’ ‘assertion,’ and ‘knowledge’. But whatever the explanation, because the truth commitment tightly intertwines with our views about belief, assertion, and knowledge, it seems best to retain that connection and revise epistemology by making compensatory

adjustments elsewhere. Once those adjustments are made, knowledge and belief turn out to be less central to epistemology than we standardly think. I do not then claim that it is epistemically acceptable to *believe* what is false or that it is linguistically acceptable to *assert* what is false. Rather, I suggest that epistemic acceptance is not restricted to belief (Cohen, 1992). Analogously, uttering or inscribing seriously and sincerely for cognitive purposes —call it ‘professing’—is not limited to asserting. Understanding is often couched in and conveyed by symbols that are not, and do not purport to be, true. Where such symbols are sentential, I call them felicitous falsehoods. I contend that we cannot understand the cognitive contributions of science, philosophy, or even our ordinary take on things, if we fail to account for such symbols. Let’s look at some cases:

Curve smoothing: Ordinarily, each data point is supposed to represent an independently ascertained truth. (The temperature at t_1 , the temperature at t_2, \dots) By interpolating between and extrapolating beyond these truths, we expect to discern the pattern they instantiate. If the curve we draw connects the data points, this is reasonable. But the data rarely fall precisely on the curve adduced to account for them. The curve then reveals a pattern that the data do not instantiate. Veritism would seem to require accepting the data only if we are convinced that they are true, and connecting these truths to adduce more general truths. Unwavering commitment to truth would seem then to require connecting all the data points no matter how convoluted the resulting curve turned out to be. This is not done. To accommodate every point would be to abandon hope of finding order in most data sets, for jagged lines and complicated curves mask underlying regularities. Nevertheless, it seems cognitively disreputable simply to let hope triumph over experience. Surely we need a better reason to skirt the data and ignore the outliers than the fact that otherwise we will not get the kind of theory we want. Nobody, after all, promised that the phenomena would accommodate themselves to the kind of theory we want.

There are often quite good reasons for thinking that the data ought not, or at least need not be taken as entirely accurate. Sometimes we recognize that our measurements are relatively crude compared with the level of precision we are looking for. Then any curve that is within some δ of the evidence counts as accommodating the evidence. Sometimes we suspect that some sort of interference throws our measurements off. Then in plotting the curve, we compensate for the alleged interference. Sometimes the measurements are in fact accurate, but the phenomena measured are complexes only some of whose aspects concern us. Then in curve smoothing we, as it were, factor out the irrelevant aspects. Sometimes we have no explanation for the data’s divergence from the smooth curve. But we may be rightly convinced that what matters is the smooth curve the data indicate, not the jagged curve they actually instantiate. Whatever the explanation, we accept the curve, taking its proximity to the data points as our justification. We understand

the phenomena as displaying the pattern the curve marks out. We thus dismiss the data's deviation from the smooth curve as negligible.

Ceteris paribus claims: Many lawlike claims in science obtain only *ceteris paribus*.³ The familiar law of gravity

$$F = Gm_1m_2/r^2$$

is not universally true, for other forces may be in play. The force between charged bodies, for example, is a resultant of electrical and gravitational forces. Nevertheless, we are not inclined to jettison the law of gravity. The complication that charge introduces just shows that the law obtains only *ceteris paribus*, and when bodies are charged, *ceteris* are not *paribus*. This is no news. ‘*Ceteris paribus*’ is Latin for ‘other things being equal’, but it is not obvious what makes for equality in a case like this. Sklar glosses it as ‘other things being normal’ (Sklar 1999, p. 702), where ‘normal’ seems to cash out as ‘typical’ or ‘usual’. Then a ‘*ceteris paribus* law’ states what usually happens. In that case, to construe the law of gravity as a *ceteris paribus* law is to contend that although there are exceptions, bodies usually attract each other in direct proportion to the product of their masses and in inverse proportion to the square of the distance between them.

This construal does not always work. Some laws do not even usually hold. The law of gravity is one. Snell's law

$$n_1 \sin i = n_2 \sin r$$

which expresses the relation between the angle of incidence and the angle of refraction of a light ray passing from one medium to another, is a second.⁴ As standardly stated, the law is perfectly general, ranging over every case of refraction. It is not true of every case, though; it obtains only where both media are optically isotropic. The law then is a *ceteris paribus* law. But it is not even usually true, since most media are optically anisotropic (Cartwright 1983, pp. 46–47). One might wonder why physicists don't simply restrict the scope of the law: ‘For any two optically isotropic media, $n_1 \sin i = n_2 \sin r$ ’. The reason is Gricean: expressly restricting the scope of the law implicates that it affords no insight into cases where the restriction does not obtain. Snell's law is more helpful. Even though the law is usually false, it is often not far from the truth. Most media are anisotropic, but lots of them—and lots of the ones physicists are interested in—are nearly isotropic. The law supplies good approximations for nearly isotropic cases. So although explanations and calculations that rely on Snell's law do not yield truths, they are often not off by much.

The law is valuable for another reason as well. Sometimes it is useful to first represent a light ray as conforming to Snell's law, and later introduce ‘corrections’ to accommodate anisotropic media. If we were interested only

in the path of a particular light ray, such a circuitous approach would be unattractive. But if we are interested in optical refraction in general, it might make sense to start with a prototypical case, and then show how anisotropy perturbs. By portraying anisotropic cases as perturbations, we point up affinities that direct comparisons would not reveal. The issue then is what sort of understanding we want. Showing how a variety of cases diverge from the prototypical case contributes valuable insights into the phenomenon we are interested in. And what makes the case prototypical is not that it usually obtains, but that it cleanly exemplifies the features we deem important.

Idealizations: Some laws never obtain. They characterize ideal cases that do not, perhaps cannot, occur in nature. The ideal gas law represents gas molecules as perfectly elastic spheres that occupy negligible space and exhibit no mutual attraction. There are no such molecules. Explanations that adduce the ideal gas law would be epistemically unacceptable if abject fidelity to truth were required. Since helium molecules are not dimensionless, mutually indifferent, elastic spheres, an account that represents them as such is false. If veritism is correct, it is epistemically unacceptable. But, at least if the explanation concerns the behavior of helium in circumstances where divergence from the ideal gas law is negligible (roughly, where temperature is high and pressure is low), scientists are apt to find it unexceptionable.

Stylized facts are close kin of *ceteris paribus* claims. They are ‘broad generalizations, true in essence, though perhaps not in detail’ (Bannock, et al. 1998, pp. 396–7). They play a major role in economics, constituting explananda that economic models are required to explain. Models of economic growth, for example, are supposed to explain the (stylized) fact that the profit rate is constant.⁵ The unvarnished fact of course is that profit rates are not constant. All sorts of non-economic factors—such as war, pestilence, drought, and political chicanery—interfere. Manifestly, stylized facts are not (what philosophers would call) facts, for the simple reason that they do not obtain. It might seem then that economics takes itself to be required to explain why known falsehoods are true. (Voodoo economics, indeed!) This cannot be correct. Rather, economics is committed to the view that the claims it recognizes as stylized facts are in the right neighborhood, and that their being in the right neighborhood is something economic models need to account for. The models may show them to be good approximations in all cases, or where deviations from the economically ideal are slight, or where economic factors dominate non-economic ones. Or the models might afford some other account of their often being nearly right. The models may differ over what is actually true, or as to where, to what degree, and why the stylized facts are as good as they are. But to fail to acknowledge the stylized facts would be to lose valuable economic information (for example, the fact that if we control for the effects of non-economic interference such as wars, epidemics, and the president for life absconding

with the national treasury, the profit rate is constant). Stylized facts figure in other social sciences as well. I suspect that under a less alarming description, they occur in the natural sciences too. The standard characterization of the pendulum, for example, strikes me as a stylized fact of physics. The motion of the pendulum that physics is supposed to explain is a motion that no actual pendulum exhibits. What such cases point to is this: The fact that a strictly false description is in the right neighborhood is sometimes integral to our understanding of a domain.

A fortiori arguments from limiting cases: Some accounts focus on a single, carefully chosen case and argue that what holds in that case holds in general. If so, it does no harm to represent the phenomena as having the features that characterize the exemplary case. Astronomy sometimes represents planets as point masses. Manifestly, they are not. But because the distance between planets is vastly greater than their size, their spatial dimensions can safely be neglected. Given the size and distribution of planets in the solar system, what holds for properly characterized point masses also holds for the planets. Another familiar example comes from Rawls. *A Theory of Justice* represents people as mutually disinterested. Rawls is under no illusion that this representation is accurate. He recognizes that people are bound to one another by ties of affection of varying degrees of strength, length, and resiliency. But, he believes, if political agents have reason to cooperate even under conditions of mutual disinterest, they will have all the more reason to cooperate when ties of affection are present. I do not want to discuss whether Rawls is right. I just want to highlight the form of his argument. If what holds for the one case holds for the others, then it does no harm to represent people as mutually disinterested. That people are mutually disinterested is far from the truth. Conceivably, no one on Earth is wholly indifferent to the fates of every other person. But if Rawls is right, the characterization's being far from the truth does not impede its function in his argument.

The foregoing examples show that in some cognitive endeavors we accept claims that we do not consider true. But we do not indiscriminately endorse falsehoods either. The question then is what makes a claim acceptable? Evidently, to accept a claim is not to take it to be true, but to take it that the claim's divergence from truth, if any, is negligible. The divergence need not be small, but whatever its magnitude, it can be safely neglected. We accept a claim, I suggest, when we consider it true enough. The success of our cognitive endeavors indicates that we are often right to do so. If so, a claim is acceptable when its divergence from truth is negligible. In that case it is true enough.

In practical, political, or prudential contexts, both the acceptance and the acceptability of falsehoods are widely recognized. One can accept, and be right to accept, the dean's latest dictum, if what matters is that the dean hath said it, not that it is true. But epistemic contexts are supposed to be

different. Many epistemologists contend that when our concerns are cognitive we should accept only what we consider true. I disagree. I suggest that *to accept* that p is to take it that p 's divergence from truth, if any, does not matter. *To cognitively accept* that p , is to take it that p 's divergence from truth, if any, does not matter cognitively. The falsehood is 'as close as one needs for the purposes at hand' (Stalnaker 1987, p. 93). In what follows, I take 'acceptance' to mean 'cognitive acceptance'.

This raises a host of issues. Perhaps the most pressing is to say something about what I mean by 'cognitive'. A familiar line is that for a consideration to be cognitive is for it to aim at truth or to be truth-conducive. Plainly, I can say no such thing. I suggest rather that a consideration is cognitive to the extent that it figures in an understanding of how things are. This is admittedly vague, but I am not sure that it is any worse than untethered remarks about truth-conduciveness and the like.

It might seem that my characterization just postpones the evil day (and not for long enough!), since 'an understanding of how things are' must itself be explicated in terms of truth or truth-conduciveness. To see the problem, compare three concepts—belief, thought, and understanding. Belief aims at truth. Roughly, a belief fulfills its goal in life only if it is true (Wedgwood, 2002; Adler, 2002). Thought, however, can be aimless. Musings, fantasies, and imaginings can be fully in order whether or not they are true. Understanding, the argument goes, is more like belief than like thought. Since there is such a thing as misunderstanding, understanding is subject to a standard of rightness. It has an aim. Misunderstanding evidently involves representing things as they are not. This suggests that the aim of understanding is truth. If so, it may seem, divergences from truth, even if unavoidable, are always cognitive defects.

The argument goes too fast. That misunderstanding involves representing things as they are not does not entail that whenever we represent things as they are not, we misunderstand them. At most it indicates that understanding is not indifferent to truth. But it does not follow that every sentence—or for that matter, any sentence—that figures in an understanding of how things are has its own truth as an objective. Understanding involves a network of commitments. It is not obvious that an aim of the network must be an aim of every, or indeed any, sentential node in the network. A goal of the whole need not be a goal of each of its parts. 'Understanding' is a cognitive success term but, in my view, not a factive. I do not expect these sketchy remarks to persuade anyone that I am right to loosen the tie between understanding and truth. My hope is that they are enough to persuade you that the jig is not yet up, that a willing suspension of disbelief is still in order.

Let us turn then to acceptance. To accept that p , I said, is to take it that p 's divergence from truth, if any, is negligible. In that case, p is true enough. Whether this is so is manifestly a contextual matter. A sentence can be true

enough in some contexts but not in others. A variety of factors contribute constraints. Background assumptions play a role. ‘A freely falling body falls at a rate of 32 ft./sec.²’ is true enough, assuming that the body is within the Earth’s gravitational field, that nothing except the Earth exerts a significant gravitational force on the body, that the effects of non-gravitational forces are insignificant, and so on. But even when these assumptions are satisfied, the formula is not always true enough, since gravity varies slightly with longitude. Sometimes it matters where in the gravitational field the freely falling body is. Whether ‘ $G = 32 \text{ ft./sec.}^2$ ’ is true enough depends on what we want the formula for, what level of precision is needed for the calculation or explanation or account it figures in. There is no saying whether a given contention is true enough independently of answering, or presupposing an answer to the question ‘True enough for what?’ So purposes contribute constraints as well. Whether a given sentence is true enough depends on what ends its acceptance is supposed to serve.

Function is critical too. If to accept that p is simply to take it that p ’s divergence from truth does not matter, it might seem that we accept all irrelevant propositions. None of my projects, cognitive or otherwise, is affected by the truth or falsity of the claim that Ethelred the Unready was a wise leader. So its divergence from truth, if any, does not matter to me. Since acceptance can be tacit, the fact that I have never considered the issue is not decisive. Nevertheless it seems wrong to say that my indifference makes the claim true enough. The reason is that the contention is idle. It performs no function in my cognitive economy. Owing to my indifference, there is no answer to the question ‘True enough for what?’

Context provides the framework. Purposes fix the ends. Function is a matter of means. The sentences that concern us tend not to have purposes or functions in isolation. Rather, they belong to and perform functions in larger bodies of discourse, such as arguments, explanations, or theories that have purposes. In accepting a sentence, then, we treat it in a given context as performing a function in a body of discourse which seeks to achieve some end. Whether ‘ $G = 32 \text{ ft./sec.}^2$ ’ is acceptable depends on whether the body of discourse it figures in serves its cognitive purpose—whether, that is, it yields the understanding of the domain that we seek.

A statement’s divergence from truth is negligible only if that divergence does not hinder its performing its cognitive function. Hence whether a contention is true enough depends not just on its having a function, but on what its function is—on what role it plays in the account it belongs to. To determine whether a statement is true enough, we thus need to identify its function. It might seem that for cognitive purposes only one function matters. If the criterion for felicity is being true enough, one might think, the function of all felicitous falsehoods is to approximate. There is, as it were, a tacit ‘more or less’ in front of all such claims. This will not do.

One reason is that the proposal is not sufficiently sensitive. Not all approximations perform the same function. Some are accepted simply because they are the best we can currently do. They are temporary expedients which we hope and expect eventually to replace with truths. We improve upon them by bringing them closer to the truth. Such approximations are, in Sellars's terms, promissory notes that remain to be discharged. The closer we get to the truth, the more of the debt is paid. They are, and are known to be, unsatisfactory. But not all approximations have this character. Some are preferable to the truths they approximate. For example, it is possible to derive a second-order partial differential equation that exactly describes fluid flow in a boundary layer. The equation, being nonlinear, does not admit of an analytic solution. We can state the equation, but cannot solve it. This is highly inconvenient. To incorporate the truth into the theory would bring a line of inquiry to a halt, saying in effect: 'Here's the equation; it is impossible to solve.' Fluid dynamicists prefer a first-order partial differential equation, which approximates the truth, but admits of an analytical solution (Morrison, 1999, pp. 56–60). The solvable equation advances understanding by providing a close enough approximation that yields numerical values that can serve as evidence for or constraints on future theorizing. The approximation then is more fruitful than the truth. There is no hope that future inquiry will remedy the situation, for it is demonstrable that the second-order equation cannot be solved numerically, while the first-order equation can.⁶ That reality forces such a choice upon us may be disappointing, but under the circumstances it does not seem intellectually disreputable to accept and prefer the tractable first-order equation. One might say that acceptance of the first-order approximation is only practical. It is preferable merely because it is more useful. This may be so, but the practice is a cognitive one. Its goal is to understand fluid flow in boundary layers. In cases like this, the practical and the theoretical inextricably intertwine. The practical value of the approximation is that it advances understanding of a domain. A felicitous falsehood thus is not always accepted only in default of the truth. Nor is its acceptance always 'second best'. It may make cognitive contributions that the unvarnished truth cannot match.

Moreover, not all felicitous falsehoods are approximations. Idealizations may be far from the truth, without thereby being epistemically inadequate. Political agents are not mutually disinterested. They are not nearly mutually disinterested. Nor is it the case that most political agents are mutually disinterested. There is no way I can see to construe Rawls's model as approximately true. Nevertheless, for Rawls's purposes, the characterization of political agents as mutually disinterested is felicitous if the features it highlights are constitutive of fair terms of cooperation underlying the basic structure of a democratic regime. There is no reason to think that in general the closer it is to the truth, the more felicitous a falsehood.

I suggest that felicitous falsehoods figure in cognitive discourse not as mistaken or inaccurate statements of fact, but as fictions. We are familiar with fictions and with reasoning about fictions. We regularly reason within the constraints dictated by a fiction, inferring, for example, that Hamlet may not have been mad, but Ophelia surely was. We also reason (with more trepidation) from a fiction to matters of fact. We may come to see a pattern in the facts through the lens that a fiction supplies. For example, we might understand the contemporary youth gang conception of being ‘dissed’ by reference to the wrath of Achilles at Agamemnon’s affront.⁷ What is needed then is an account of how fictions can advance understanding.

David Lewis (1983) interprets fictional statements as descriptions of other possible worlds. We understand them in the same way that we understand ordinary statements of fact. For example, we understand the Sherlock Holmes stories in the same way we understand histories of Victorian England, the crucial difference being that the histories pertain to the actual world, while the stories pertain to other possible worlds. Lewis’s realism about possible worlds is hard to countenance, but we need not enter into debates about it here. For even if we accept his metaphysics, a problem remains. It is puzzling how knowing what happens in another possible world can afford any insight into what happens in this world. Lewis measures the proximity of possible worlds in terms of similarity (Lewis 1973). The closest possible worlds are the ones most similar to the actual world. But even if one thinks that there is a non-question-begging way to assess the relative similarity of different possible worlds, the gap remains. It is hard to see how the knowledge that a nearby world is populated by rational economic agents should contribute to our understanding of the economic behavior of actual human agents.

Kendall Walton (1990) construes fiction as make-believe. To understand a fiction is to make believe or pretend that it is true. Stephen Yablo uses Walton’s account to underwrite fictionalism in the philosophy of mathematics (1998, 2002). We can, Yablo maintains, avoid ontological commitment to mathematical entities by construing mathematics as an elaborate, highly systematic fiction. Then in doing mathematics, we make believe that there are such entities and make believe that they are related to one another in the ways the theorems say. Although at first glance this seems promising, it faces two serious problems. First, what it is to pretend is by no means clear. Exactly how one can pretend that power sets or cube roots exist if they do not, is not obvious. Second, the problem we saw with Lewis’s theory also plagues Walton and Yablo. It is not at all clear how a pretense illuminates reality—how, for example, pretending that human beings are rational economic agents provides any insight into actual human economic behavior.

Philip Kitcher (1993) suggests that scientific idealizations are stories whose referents are fixed by stipulation. The scientific statements involving the idealizations are true by convention. The gap remains. If a story is to

advance scientific understanding we need to know the connection between the story and the facts. Otherwise, the realization that certain relations among pressure, temperature, and volume are true by stipulation in the story of the ideal gas leaves us in the dark about the relations among pressure, temperature, and volume in actual gases.

The gap can be bridged by appeal to exemplification, the device by which samples and examples highlight, exhibit, display, or otherwise make manifest some of their features (Goodman 1968, pp. 45–67). To make this out requires a brief discussion of the device. A commercial paint sample consists of a patch of color on a card. The patch is not merely an instance of the color, but a telling instance—an instance that exemplifies the color. By so doing, the sample equips us to recognize the color, to differentiate it from similar shades. The sample then affords epistemic access to the color. Although the patch on the sample card has a host of other features—size, shape, location, and so on—it standardly does not exemplify them. Exemplification is selective. It brings out some features of an exemplar by overshadowing, downplaying, or marginalizing others. Nothing in the nature of things dictates that the patch's color is worthy of selection, but its shape is not. What, if anything, an item exemplifies depends on its function. The very same item might perform any of a variety of functions. The patch on the sample card could be used to teach children what a rectangle is. In that case, it would exemplify its shape, not its color. The sample card could be used as a fan. Then the patch would not exemplify at all. Exemplification is not restricted to commercial and pedagogical contexts. Whatever an item exhibits, highlights, or displays, it exemplifies. A poem might exemplify its rhyme scheme, its imagery, or its style. A water sample might exemplify its mineral content, its flavor, or its impurities. Exemplification, I have argued elsewhere, is ubiquitous in art and science (Elgin 1996, pp. 171–183).

Treating paint samples as paradigmatic exemplars may encourage the idea that exemplified features are all like expanses of color—homogeneous qualities spread out before us, lacking depth and complexity, hence able to be taken in at a glance. Many are not like that. Pick up a rock containing iron ore. It might serve as a sample of iron, or of hematite, or of something that bears a striking resemblance to your high school algebra teacher. It can exemplify such features only where certain background assumptions are in place. Not just anyone looking at the rock could tell that it exemplified these features.

Moreover, although in principle any item can serve as an exemplar and any feature can be exemplified, a good deal of effort may be required to bring about the exemplification of a recondite feature. Some of that effort is mental. Just as we ignore the shape of the paint sample and focus on the color, we can ignore the fact that the rock looks like your algebra teacher and focus on its hardness. This is a start. But some irrelevant features so intricately intertwine with relevant ones that more drastic measures are

called for. If we seek to exemplify recondite features of iron, mental agility alone may not be enough to bracket the effects of other minerals in the rock. So we refine the ore and filter out the impurities. The result of our efforts is pure iron. It is a product of a good deal of processing⁸ which eliminates complicating factors and brings to the fore characteristics that are hard to detect and difficult to measure in nature. To facilitate the exemplification of the feature of interest, we do not just mentally sideline features we consider irrelevant, we physically remove some of them.

Even then, we do not just contemplate the bit of iron as we might a paint sample. We subject it to a variety of tests. We seek to produce circumstances where the features of interest stand out. We not only investigate the iron's behavior in standard conditions, we study what happens in extreme conditions—very high or low pressure or temperature, in a vacuum, under intense radiation, and so forth. Although we recognize that the test conditions do not ordinarily (or perhaps ever) obtain in nature, we take it that the behavior of the refined metal in the test conditions discloses something about the natural order. If so, by understanding what happens in the lab, we can understand something of what happens in the world. The connection is, of course, indirect. It involves a complicated extrapolation from situations and materials that are highly artificial and carefully contrived. One might argue that the lab itself is a fictional setting and the conclusions we draw about nature on the basis of our laboratory findings are projections from fiction to fact. I don't quite want to say that (although I suspect that Nancy Cartwright does). But I do want to point out that experimentation involves a lot of stage setting.

There is a tendency to think of experiments as processes that generate information, hence as ways to find things out. This of course is true. But it is worth noting that an experiment is not like an oracle,⁹ or an anchorman, or a fortune cookie. It does not just issue a report stating its results. It displays them. It shows what happens to the magnetic properties of iron in conditions near the melting point. The experiment exemplifies its results.

No matter how carefully we set the stage, irrelevancies remain. We do not and ought not read every aspect of the experimental result back onto the world. Not only are there irrelevant features, there are issues about the appropriate vocabulary and level of precision for characterizing what occurs. The fact that the experiment occurred in Cleveland is unimportant. The fact that the sample has a certain mass or lattice structure may or may not be significant. The fact that the temperature is above 770° C may matter, while the fact that it was above 790° C does not. Some features of the iron in the experimental situation are telling features. Others are not. The telling features are the ones that the experiment discloses or makes manifest. By exemplifying certain features, the experiment brings them to light and affords epistemic access to them. That is its cognitive contribution. Other features, though equally real, are not exemplified. The experiment

embodies an understanding of the phenomenon in question through its exemplification of telling features. By making these features manifest, it affords an understanding of the phenomenon.

If the cognitive contribution of an exemplar consists in the exemplification of select features, then anything that exemplifies exactly those features can, in a suitable context, make the same contribution. Return for a moment to the paint sample. I spoke of it as though it is a sample of *paint*, a telling instance of the stuff you might use to paint the porch. This is not true. The sample on the card does not consist of paint, but of an ink or dye of the same color as the paint whose color it exemplifies. If the sample were supposed to exemplify the paint itself, or the chemical features of the paint, the fact that it is not paint or has a different chemical composition would be objectionable. But since it exemplifies only the color, all that is needed is something that is the same color as the paint. The exemplar need not itself be paint. Similarly in scientific cases. Consider a DNA molecule that exemplifies its molecular structure. Anything that exemplifies the same structure has the capacity to perform the same function in our understanding of DNA. No more than the paint sample needs to consist of paint, does the exemplar of DNA's molecular structure need to consist of DNA. A schematic model that exemplifies the same features but has a different material (or even immaterial) substrate could do the job.

Here is where felicitous falsehoods enter the picture. Something other than paint can serve as a paint sample, affording epistemic access to a color also instantiated by the paint. Something other than a molecule can exemplify molecular structure, thereby affording epistemic access to a structure also instantiated by the molecule. A felicitous falsehood then is a fiction that exemplifies a feature in a context where the exemplification of that feature contributes to understanding. The utility of such a falsehood is plain. It is sometimes inconvenient, difficult, or even impossible to bring it about that the phenomena exemplify all and only the features that interest us. (DNA molecules are very small, charged pions are short lived.) If we introduce a falsehood that exemplifies those features—a bigger, longer-lasting model, for example—we can highlight them and display their significance for the understanding of the phenomenon in question. By exemplifying features it shares with the phenomena, a felicitous falsehood affords epistemic access to how things actually are. The camel's nose is now officially inside the tent.

There is more than one role that such fictions can play. Some serve as points of reference. We understand things in terms of them. In the simplest cases, like the model displaying the helical structure of the DNA molecule, they are simply schemata that exemplify factors they share with the phenomena they concern. They qualify as fictions because they diverge from the phenomena in unexemplified properties. (DNA molecules are not made of tinker toys.) In other cases, the connection to the facts is less direct. No real

gas has the properties of the ideal gas. The model is illuminating though, because we understand the properties of real gases in terms of their deviation from the ideal. In such cases, understanding involves a pattern of schema and correction. We represent the phenomena with a schematic model, and introduce corrections as needed to closer accord with the facts. Different corrections are needed to accord with the behavior of different gases. The fictional ideal then serves as a sort of least common denominator that facilitates reasoning about and comparison of actual gases. We ‘solve for’ the simple case first, then introduce complications as needed.

Acknowledging the role of corrections might seem to suggest that the detour through fictions is just a circuitous route to the truth. Rather than a simple true description of the behavior of neon, we get a complicated truth that makes reference to deviations from some ideal. But the full cognitive contribution of the exercise resides in the truth, not the fiction. I don’t think this is right for several reasons. The first is that sometimes the corrections that would be needed to yield a truth are unnecessary or even counter-productive. A *fortiori* arguments from limiting cases succeed because the corrections are unnecessary. If a consideration holds for one case, even if that case is a fiction, it holds for all. The multiple and complicated ways actual cases diverge from the fictive ideal make no difference. If Rawls’s argument is sound, ‘correcting for’ ties of affection just muddies the waters. Nor is this the only case where fidelity to the facts can prove a hindrance. It follows from the ideal gas law that as volume goes to zero, pressure becomes infinite. This would not happen. Given a fixed number of molecules, pressure increases as volume decreases—not to infinity, but only to the point where the container explodes. No one of course denies this. But to understand what would happen in the limit, we need to prescind from such material inconveniences and pretend that the walls of the container are infinitely strong. We need then to introduce not corrections that bring us back to the facts, but further idealizations. A second reason is that the requisite corrections often yield not truths, but more refined models. They supplant one falsehood with another. A third reason is that even where corrections yield truths, the fiction may be more than a *façon de parler*. It can structure our understanding in a way that makes available information we would not otherwise have access to. If, e.g., we draw a smooth curve that skirts the data, and construe the data as a complex of relevant and irrelevant factors (signal and noise), or construe a transaction in terms of an economic model overlaid with non-economic factors which skew the outcome, we impose an order on things, highlight certain aspects of the phenomena, reveal connections, patterns and discrepancies, and make possible insights that we could not otherwise obtain (Dennett 1991). We put ourselves in a position to see affinities between disparate occurrences by recognizing them as variations on a common theme.

Still, to say that felicitous falsehoods figure ineliminably in our understanding of certain phenomena is not to say that without them we would have no understanding whatsoever of those phenomena. Even without Snell's law, inspection and instantial induction would provide some insight into what happens when light passes between air and water, or between any other two specified media. What we would lack is a systematic understanding of how the several cases are alike. Felicitous falsehoods configure a domain, enabling us to characterize the phenomena in ways that would otherwise be unavailable. When, for example, we can construe seemingly divergent phenomena as variants of a common scheme, or as perturbations of a regular pattern, or as deviations from a simple norm, we see them and their relations to one another in a new light. We can discern systematic interconnections that direct inspection of the facts would not reveal. The fictions and the configurations of the domains that they engender provide conceptual resources for representing and reasoning about the phenomena in new and sometimes fruitful ways.

If I am right, not every theory is a conjunction of sentences, all of which are supposed to be true. Rather, a theory may be composed of both factual and fictional sentences, and the fictional sentences may play any of several different roles. This means that to understand what a theory conveys, and to understand the phenomena in terms of the theory requires sensitivity to the different roles the different sentences play. And to assess a theory requires determining whether the component sentences are true enough for the parts they are assigned to play. Yablo is investigating the metaphysical implications of the idea that fictions infuse even our most fundamental theories (Yablo 1998; 2002). I want to underscore the epistemological implications. One obvious consequence is that it is not plausible to think of an acceptable theory as a mirror of nature. Even if the goal of a theory is to afford an understanding of a range of facts, it need not approach or achieve that goal by providing a direct reflection of those facts.

I have argued that a variety of components of cognitively acceptable theories neither are nor purport to be true. Rather, they are fictions that shed light on the phenomena they concern. They thereby contribute to our understanding of those phenomena. Even if my account makes sense of models and idealizations in science, a worry remains. My position threatens to make the world safe for postmodernist claptrap. If truth is not required for epistemic acceptability, why isn't a flagrantly false account acceptable? What is the objection to claiming that a theory attesting to the healing powers of crystals is as acceptable as the theories constituting mainstream crystallography? We seem to lose a valuable resource if we can't simply say, 'Because it is false!' My seemingly wimpy requirement that an acceptable account must yield an understanding of how things are gives us what we need. An account that yields such an understanding must accommodate the facts in a domain. The accommodation may be indirect. Strictly false

idealizations may be deployed. Detours through stylized facts may be made. The justification for the falsehoods is that they figure in accounts that make sense of the facts. A cognitively acceptable account sheds light on its subject. Where felicitous falsehoods are involved, the light may be oblique.

A theory can claim to make sense of a range of facts only if it is factually defeasible—only if, that is, there is some reasonably determinate, epistemically accessible factual arrangement which, if it were found to obtain, would discredit the theory. A felicitous falsehood is acceptable only if the theory or system of thought it belongs to accommodates the epistemically accessible facts. Exactly what this requires needs to be spelled out. The usual considerations about evidence, simplicity, scope, and so forth come into play. Even though some of the sentences in a theory are not supposed to be true, the way the world is constrains the acceptability of the theory they figure in. If, for example, evidence shows that friction plays a major role in collisions between gas molecules, then unless compensating adjustments are made elsewhere, theories that model collisions as perfectly elastic spheres will be discredited. An acceptable theory must be at least as good as any available alternative, when judged in terms of currently available standards of cognitive goodness. So such a theory would also be discredited by a theory that better satisfied those standards. Neither a defeated nor an indefeasible theory is tenable. Because it is indifferent to evidence, claptrap is indefeasible. Hence it is untenable. I said earlier that even in theories that include felicitous falsehoods truth plays a role. We see now what the role is. A factually defeasible theory has epistemically accessible implications which, if found to be false, discredit the theory. So a defeasible theory, by preserving a commitment to testable consequences retains a commitment to truth. My account then does not turn science into a genre of fiction. Quine said that the sentences of a theory face the tribunal of experience as a corporate body (Quine, 1961, p. 41). Although he recognized that not all the members of the corporation could be separately tested against experience, he probably believed that all were supposed to be fine, upstanding truths. I suggest that this is not so. An acceptable theory must make its case before the tribunal of experience. But the members of the corporation are more various and some a bit shadier than Quine suspected.

Notes

* I would like to thank Jonathan Adler, Nancy Nersessian, Amelie Rorty, Robert Schwartz, Bas van Fraassen, Stephen Yablo, and members of my audiences at Bryn Mawr College, Tufts University, and the 2003 Rutgers Epistemology Conference for helpful comments on earlier drafts of this paper.

1. I use the term ‘theory’ broadly, to comprehend the ways mature sciences and other disciplines account for the phenomena in their domains. Under my usage, a physical theory would include its models rather than being distinct from them.
2. I do not have strong intuitions about this case, but I do not think it is clearly wrong to say that the cat has such a belief.
3. Whether so-called *ceteris paribus* laws are really laws is a subject of controversy. See, for example, *Erkenntnis* 57, 2002, for a range of papers on the issue. Although for brevity I speak of them as laws, for my purposes, nothing hangs on whether the generalizations in question really are laws, at least insofar as this is an ontological question. I am interested in the role such generalizations play in ongoing science. Whether or not they can (in some sense of ‘can’) be replaced by generalizations where all the caveats and restrictions are spelled out, in practice scientists typically make no effort to do so. Nor, often, do they know (or care) how to do so.
4. i and r are the angles made by the incident beam to the normal and n_1 and n_2 are the refractive indices of the two media.
5. ‘The profit rate is the level of profits in the economy relative to the value of capital stock’ (Bannock et al. 1998, p. 397).
6. Conceivably, of course, the equations in question will be superceded by some other understanding of the subject, but the fact that the equation we consider true does not have an analytical solution provides no reason to think so. Nor does it provide reason to think that the considerations that supersede it will be mathematically more tractable, much less that in the long run science will be free of all such irksome equations.
7. I am indebted to Amelie Rorty for this example.
8. This is why Nancy Cartwright thinks the laws of physics lie. The laws are developed on the basis of, and are strictly true only of, the processed samples, not their naturally occurring counterparts.
9. Actually, of course, I don’t know how oracles are supposed to operate. I have always assumed that they simply emit true sentences like ‘Socrates is the wisest of men.’

References

- Adler, Jonathan. (2002) *Belief's Own Ethics*, Cambridge, MA: MIT.
- Bannock, Graham, R.E. Baxter, Evan Davis. (1998) *Dictionary of Economics*, New York: John Wiley and Sons.
- Bonjour, Laurence. (1985) *The Structure of Empirical Knowledge*, Cambridge, MA: Harvard.
- Cartwright, Nancy. (1983) *How the Laws of Physics Lie*, Oxford: Oxford.
- Cohen, L. Jonathan. (1992) *An Essay on Belief and Acceptance*, Oxford: Clarendon.
- Dennett, Daniel. (1991) ‘Real Patterns,’ *Journal of Philosophy* 88: 27–51.
- Elgin, Catherine Z. (1996) *Considered Judgment*, Princeton: Princeton.
- Goldman, Alvin. (1999) *Knowledge in a Social World*, Oxford: Clarendon.
- Goodman, Nelson. (1968) *Languages of Art*, Indianapolis: Hackett.
- Kitcher, Philip. (1993) *The Advancement of Science*, Oxford: Oxford.
- Lehrer, Keith. (1986) ‘The Coherence Theory of Knowledge,’ *Philosophical Topics* 14: 5–25.
- Lewis, David. (1983) ‘Truth in Fiction,’ *Philosophical Papers I*, Oxford: Oxford, pp. 261–281.
- . (1973) *Counterfactuals*, Cambridge, MA: Harvard.

- Morrison, Margaret. (1999) 'Models as Autonomous Agents,' *Models as Mediators*, ed. Mary S. Morgan and Margaret Morrison, Cambridge: Cambridge, pp. 38–65.
- Quine, W. V. (1961) 'Two Dogmas of Empiricism,' *From a Logical Point of View*, New York: Harper & Row, pp. 20–46.
- Sklar, Lawrence. (1999) 'Philosophy of Science,' *Cambridge Dictionary of Philosophy*, ed. Robert Audi, Cambridge: Cambridge, pp. 700–704.
- Sosa, Ernest. (2000) 'Skepticism and Contextualism,' *Skepticism*, ed. Ernest Sosa and Enrique Villanueva, Boston: Blackwell, pp. 1–18.
- Stalnaker, Robert. (1987) *Inquiry*, Cambridge, MA: MIT Press.
- Walton, Kendall. (1990) *Mimesis as Make-Believe*, Cambridge, MA: Harvard.
- Wedgwood, Ralph. (2002) 'The Aim of Belief,' *Philosophical Perspectives* 16: 267–297.
- Yablo, Stephen. (2002) 'Abstract Objects,' *Philosophical Issues* 12: 220–240.
- . (1998) 'Does Ontology Rest on a Mistake?' *Proceedings of the Aristotelian Society*, Supplementary Volume 72: 229–261.