THREE WAYS OF GETTING IT WRONG: INDUCTION IN WONDERLAND

Brendan Shea, PhD, University of Illinois

Alice encounters at least three problems in her struggles to understand and navigate Wonderland. The first arises when she attempts to predict what will happen in Wonderland based on what she has experienced outside of Wonderland. In many cases, this proves difficult—she fails to predict that babies might turn into pigs or that a grin could survive without a cat. Alice's second problem involves her efforts to figure out the basic nature of Wonderland. So, for example, it is difficult for Alice to see how she could *prove* that her experiences were the result of her dreaming and not something else. The final problem is manifested by Alice's attempts to understand what the various residents of Wonderland mean when they speak to her. In Wonderland, "mock turtles" are real creatures and people go places with a "porpoise" (and not a purpose). All three of these problems concern Alice's attempts to infer information about unobserved events or objects from those she has observed. In philosophical terms, they all involve *induction*.

Induction, it turns out, is hugely important to our daily lives. It is inductive reasoning that allows us to figure out which foods nourish and which poison; induction guides our conviction of criminals and pardoning of the innocent; induction is the basis of all scientific knowledge. Induction also provides our sole basis for understanding *language*. The problem with induction, as Alice learns, is that one always risks being wrong. In Wonderland, many of Alice's inductively supported beliefs turn out suddenly to be false, and she is forced to start from scratch. Her successes and failures in doing so offer a window into the working of inductive reasoning, and should therefore be of real interest to us.

BELIEVING IMPOSSIBLE THINGS: INDUCTION AND DEDUCTION

Induction, as it usually understood by philosophers, is a type of reasoning that stands in explicit contrast to *deduction*. In order to reason deductively from known facts to a conclusion, one must show that it is *impossible* for the conclusion to be false on the assumption that the original facts are as we thought. Lewis Carroll offers the following example of a deductive argument in his *Symbolic Logic*: "All cats understand French; some chickens are cats" therefore, "Some chickens understand French."[1] The conclusion is, of course, false. However, if one *assumes* that the first sentence is true, there is *no way* the second sentence could be false. In using deduction, one never really goes beyond what one knows—one merely restates it in new (and sometimes interesting) ways.

In contrast to deductive reasoning, inductive reasoning attempts to go beyond what is already known. So, for example, were Alice to reason from *No rabbits have ever spoken to me before today* to the conclusion that *No rabbits will speak to me today* she would be reasoning inductively. This is because, as Alice discovers, it is *possible* that the first sentence could be true and the second false.

Most, if not all, of all our everyday beliefs about the world have been arrived at by inductive reasoning. Alice provides a few good examples of inductively supported beliefs when she considers drinking the bottle labeled DRINK ME:

It was all very well to say "Drink me" but the wise little Alice was not going to do that in a hurry. "No, I'll look first," she said, "and see whether its marked 'poison' or not"; for she had read several nice stories about children who got burnt, and eaten up by wild beasts, and other unpleasant things, all because they would not remember the simple rules their friends had taught them: such as, that a red-hot poker will burn you if you hold it too long; and that, if you cut your finger very deeply it usually bleeds; and she had never forgotten that, if you drink much from a bottle marked 'poison,' it is almost certain to disagree with you, sooner or later.[2]

All of the rules Alice cites are good (if simple) ones. Nevertheless, we have only inductive reasoning to support them. For instance, the reason we think that fires will burn us in the future is because we know that they have burned us in the past. The same holds true any time we believe something on the authority of some reliable teacher, parent or book. In fact, nearly all of our basic beliefs about history, science or other people have been arrived at by inductive reasoning.

DOWN THE RABBIT HOLE: HUME'S PROBLEM WITH PREDICTION

One common type of inductive reasoning involves our attempts to *predict* what we will experience in the future based on what has happened in the past. For example, most of us believe that rabbits will

not begin talking nor babies suddenly become pigs. But suppose that somebody (we'll pretend it is Humpty Dumpty) disagrees with us about these beliefs—Humpty claims that, at 6:00am tomorrow morning, the world we know will suddenly turn into Wonderland. Animals will talk, playing cards will hold trials, and people will change size dramatically when they eat food. What, if anything, could we say to convince Humpty that he is wrong?

In his famous book *An Enquiry Concerning Human Understanding*, David Hume (1711-1776) argues that there is nothing we could say to rationally persuade someone like Humpty that he is wrong. Hume's argument proceeds as follows. First, he divides the types of knowledge one can have into two categories: that pertaining to *Relations of Ideas* and that pertaining to *Matters of Fact*.[3] Hume thinks that knowledge of the former category is possible, but it pertains only to truths we can figure by reflecting on the nature of our own ideas. For example, Hume thinks we can know that all triangles have three sides and that all bachelors are unmarried, but this knowledge doesn't depend on the *existence* of any triangles or bachelors. Even if every male in the world were married, for example, it would remain true that *if* there were a bachelor, he would be unmarried. This is guaranteed by our idea of bachelor-hood. Hume might be quite happy with Humpty's observation that we are the "master" of our words.[4] According to Humpty's story, we can figure out whether "all triangles have three sides" is true or false simply by considering what we mean by words like 'triangle.'

In addition to our knowledge of Relations of Ideas, Hume grants that we can have knowledge about the Matters of Fact we have already *observed*. For example, we know that we have not up until now observed rabbits talking or playing cards holding trials. Hume's problem of prediction concerns the possibility of knowing about *unobserved* Matters of Fact. Humpty's challenge is just a more specific version of this problem. We have not yet observed what will happen at 6:00am tomorrow; Humpty challenges us to justify our belief the world will not suddenly change into Wonderland.

The difficulty of Hume's problem becomes apparent when we consider how we might do this. When we defend claims about Relations of Ideas, we can appeal to the unimaginability of its being false for our justification. So, for instance, it is impossible to imagine a non-three-sided triangle or a female bachelor. But Humpty's claim that the world will become Wonderland isn't impossible to imagine; in fact, we visualize Wonderland whenever we read Carroll's work. As Hume notes, we can imagine *any* matter of fact being different: "*That the sun will not rise tomorrow* is no less intelligible a proposition, and implies no more contradiction, than the affirmation, *that it will rise*."[5] Nor can we

defend our prediction as we do our beliefs about Matters of Fact we have already observed. Our beliefs about observed Matters of Fact are based on what we have seen, smelt, touched, and tasted. We can't sense the future in this way.

According to Hume, the only reason we don't think that the world will radically change tomorrow is that it hasn't ever changed in this way *before*. In fact, Hume thinks all of our beliefs about unobserved matters rest on one key assumption—*that the future will resemble the past*. The problem is that this assumption is *itself* a belief about an unobserved matter of fact. Moreover, it is precisely what Humpty was asking us to defend. He agrees that, up until today, the world has never radically changed. He just thinks that, starting tomorrow, it will. And it seems we have no rational way of convincing him he is wrong.

Hume's conclusion is not that we have no basis for making predictions, but rather that our ability to do so successfully is quite independent of our deductive reasoning ability. Alice, for instance, shows little evidence of being an expert on deductive logic. However, she uses inductive reasoning with great success. In learning how to change her size by the consumption of various foods and drinks, for example, Alice is using inductive reasoning in order to make successful predictions. Her use of evidence about past events to predict and control the future course of nature is prototypical of "scientific" reasoning, and gives some idea of just how important prediction is to our everyday lives. Humpty Dumpty, who by contrast seems fairly competent at deductive logic, provides a good example of a poor inductive reasoner. When Alice first encounters Humpty, he is singing a song about how all the King's men won't be able to put him back together again. When queried by Alice, however, Humpty seems oblivious to the obvious predictive relevance of such a song, and refuses to move from his precarious perch. Humpty, despite his argumentative acumen, seems destined for a poor end.

WHOSE DREAM IS THIS, ANYWAY? THE UNDERDETERMINATION OF THEORY BY EVIDENCE

From now on, let's suppose we've solved Hume's problem, and have managed to become perfect predictors. That is, let's suppose we're perfect at predicting what will happen to *us* in the future– what we will see, touch, taste and smell for the rest of our lives. We now come to a new problem concerning induction involving our ability to use this knowledge to justify beliefs about things we

can't sense in this way. Consider Alice's dilemma at the end of *Through the Looking Glass*, when she is trying to determine whether her experiences were the result of *her* dream or whether they were the result of the *Red King's* dream. It seems obvious to us, of course, that it must have been Alice's dream. The problem for Alice, however, is to explain to a skeptic (like Tweedledee) *why* she is justified in believing that it is her dream.

Alice's problem is a special case of what philosopher W.V. Quine (1908-2000) calls the problem of *underdetermination* of theory by evidence[6]. Here, a person's *theory* is simply the collection of all of her beliefs about the world. Alice's theory, for instance, includes such beliefs as *London is located in England* and *I am not a character in the Red King's dream*. Tweedledee's competing theory, by contrast, includes the belief that *Alice is a character in the Red King's dream*. Someone's *evidence*, in Quine's sense, consists of everything that she can sense. Alice's evidence, for example, includes her memory of seeing the Red King sleeping and her memory of apparently "waking up." The problem, according to Quine, is that no matter how much evidence we gather, there will always be multiple incompatible theories that can explain all of it.

Consider Alice's experience of "waking up," for instance. This might seem to rule out Tweedledee's theory. After all, if Alice woke up, it must have been her dream, right? Tweedledee has an easy response, however—the reason that it seems to Alice that she is real and has woken up is because this is exactly what the Red King is dreaming. Tweedledee makes just this point when refuting Alice's claim that her tears show her reality: "I hope you don't suppose those are *real* tears."[7] In fact, it turns out there is no possible piece of evidence that Tweedledee cannot accommodate by making suitable changes elsewhere in his theory. On the basis of such considerations, Quine concludes that nearly any statement can be reconciled with any piece of evidence: "any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system."[8] No amount of evidence-gathering, it appears, will allow Alice to prove Tweedledee wrong.

The relevance of underdetermination to everyday life becomes apparent when one notes that Tweedledee's style of argument can be applied to challenge *any* belief we have about things that can't be directly sensed. And we have lots of such beliefs. Many scientific theories, for example, posit the existence of things too small or strange to be directly sensed: quarks, gamma rays, electrons, and gravitational forces. Scientists, using inductive reasoning, believe that these things exist because their existence explains the types of things we can examine with our senses. The problem of

underdetermination states that there will always be some *other* theory (incompatible with ours) that could *also* explain this evidence.

It is not just scientists who believe in things they can't see. Alice, for example, believes that drinking the potion caused her to shrink and not the rest of the world to grow. However, there is nothing she could sense that would allow her to determine which process is actually happening. Alice, like the rest of us, also believes in object *permanence*. So, for example, she believes that the Duchess she encounters at the Queen's party is the *same* Duchess she saw earlier. But this isn't the only theory that could explain Alice's experience. One might alternatively think that the world is being destroyed and instantaneously created anew each second. We don't remember any such destructions, of course, but that is only because we were created so as to have (false) memories of a continuous experience. According to this theory, Alice is in fact seeing a different Duchess than the one she remembers meeting. Once again, there is nothing Alice could ever do to discover whether such a theory was true or false.

Like Hume, Quine does not intend for his problem to induce skepticism about the value and legitimacy of scientific inquiry or our ordinary ways of doing things. Instead, he is merely pointing out that successful inductive reasoning requires more than 1) deductive reasoning or 2) making successful predictions. Quine goes on to suggests three criteria by which we decide which theories we adopt.[9] First, we need to make sure our theory doesn't contain false claims about our sensory experiences. For example, Alice shouldn't adopt the theory that she's never had a Wonderland-experience. Second, we should adopt theories that are as *simple* as possible. It's simpler, all things considered, to believe that there was only was only *one* Duchess instead of two. Finally, when we add new beliefs to our theory, we ought to change as *few* of our old beliefs as possible. If Alice were to believe that she were a character in the Red King's dream, for instance, she would have to change nearly every other belief she had (all of which presumed that she was real). The belief that the land beyond the looking glass was her own dream, in comparison, fits quite well with the rest of Alice's beliefs.

In the end, of course, there is no way of *ensuring* that one's theory is correct, just as there is no system for making perfect predictions. It might simply turn out that both Alice and Tweedledee, even after comparing evidence, are both justified in their respective beliefs, despite these beliefs contradicting one another. The correct response to this, according to Quine, is simply to note that

that one *must* continue to believe in the truth of *some* theory; without it, one couldn't get around in the world. Alice's solution of recognizing the problem and simply going about her life is thus perfectly fine.

THERE'S NO MEANING IN IT, AFTER ALL: THE INDETERMINACY OF MEANING

Before moving onto the final problem, let's pretend that we've solved the first two. We're perfect at predicting what will happen to us, and we've come up with a complete, true theory of how the physical world works. The last problem of induction asks us to take this knowledge and use it to determine what various people *mean* by their words. This is a source of constant frustration to Alice, who has to deal with such nonsensical characters as the Mad Hatter, the Queen of Hearts, and Humpty Dumpty. Almost every character in Wonderland and beyond the looking glass misuses or equivocates on some key English word or phrase, and Alice is tasked with using her inductive reasoning skills to figure out what is meant *in English* by Wonderland words such as 'Time,' 'whiting,' and so on.

Quine takes up the problem of determining what people mean by their words in a book called *Word and Object*. In order to simplify matters, he supposes that a hypothetical translator finds himself (like Alice) alone in an environment in which it is unclear what various people mean by their words. The translator has no dictionary and no access to any bilinguals who speak both English and the native language (which I'll call Wonderese). After a careful consideration of the process by which this translation might proceed, Quine concludes that there will always be multiple incompatible ways to translate Wonderese words into English–that is, it is *indeterminate* what people mean by their words. To see how he arrives at this strange conclusion, it will help to consider Alice's struggles in a little more detail.

The first step in any translation is to establish which native sounds mean *yes* and *no*. Alice herself notes the reason for this when she fails in her attempt to translate her kittens' sounds into English:

It is a very inconvenient habit of kittens (Alice had once made the remark) that, whatever you say to them, they always purr. 'If they would only purr for 'yes' and mew for 'no,' or any rule of that sort' she had said, 'so that one could keep up a conversation! But how can you talk with a person if they always say the same thing?[10]

Once we have figured out which native sounds mean *yes* and *no*, we can ask the natives questions to determine what they mean by various other sounds. It is through this process, for instance, that Alice discovers that 'whiting' is the stuff used to polish shoes beneath the ocean and that 'Time' is somebody that the Mad Hatter knows.

The problem arises, according to Quine, when one realizes that there will always be multiple English translations that can explain the answers Wonderland residents give to Alice's questions. So, suppose that Alice is trying to figure out whether the Wonderese speakers mean the same thing she does by 'rabbit.' She first points to the White Rabbit and asks a group of them 'Rabbit?' They respond affirmatively. In order to be extra careful, Alice points to the White Rabbit and his rabbit relatives at various times and places, and under all sorts of conditions. The Wonderland inhabitants continue to agree. However, Alice cannot establish that the Wonderese word 'rabbit' means the same thing as the English word until she rules out the possibility that the Wonderland inhabitants are referring to something that is *co-located* with rabbits. Some things that are co-located with the White Rabbit, for example, are all of his various *parts* (arms, legs, and so on). Perhaps the Wonderese word 'rabbit' refers only to some essential part of a rabbit (the heart, for example) when this part is attached to the rest of the rabbit. It is in this spirit that Quine writes "Point to a rabbit and you have pointed to a stage of a rabbit, to an integral part of a rabbit, to the rabbit fusion, and to where rabbithood is manifested."[11]

The problem becomes even worse when one is trying to figure out the meanings of abstract words such as 'intelligence,' 'courage,' or 'acceleration.' The things that these words refer to are quite difficult (or impossible) to point to. The Dormouse's question "did you ever see such thing as a drawing of a muchness!"[12] might strike us as quite silly, but it presents a significant problem for Quine's translator. In the case of sayings like 'much of a muchness,' there might be quite a number of sayings in other languages that could serve as adequate "translations."

As was the case with the underdetermination of theories, there are some limits on the translations we can adopt. In particular, we shouldn't adopt translations that get the natives' patterns of assent and dissent wrong. For example, Alice has good reasons not to translate the Wonderese 'rabbit' as meaning *mouse*. The problem is that, once we have satisfied this minimal constraint, we have no grounds for convincing someone who has a *different* translation that she ought to adopt ours, instead. Suppose, for instance, that the Red Queen were to visit Wonderland with Alice. She would

presumably try to come up with a translation that made the connections to her own language as simple (Looking-Glass-ish) and straightforward as possible, while Alice would try to do the same with English. In the end, the translations may disagree, even after she and Alice have carefully settled all the facts not related to meaning.

In the real world, there are a variety of factors that keep us from disagreeing on translations of particular languages. It is very difficult, after all, to inductively reason our way to a translation based solely on our observations of native speakers. It is much easier to rely on dictionaries or textbooks, many of which are based on decades of careful research by linguists. In Wonderland, by contrast, this indeterminacy presents a much more obstinate and obvious problem. Wonderland residents repeatedly misunderstand both Alice and each other in strange (and sometimes malicious) ways. When Alice objects that the note in the knave's trial means nothing, for example, the king exclaims "If there's no meaning in it that saves a world of trouble, you know, as we needn't try to find any"[13] before launching into his own biased translation. Humpty Dumpty seems to have much the same philosophy when he offers Alice a translation of "Jabberwocky."

Not surprisingly, Alice's actions again suggest the best path for us to follow. In most cases, she strives to translate people's words in the way that she would want her own words understood. If it seems crazy to Alice that people would mean a certain thing by a particular word, she usually assumes that they *don't* mean that until their actions prove her wrong. Her attempts to provide translations are thus based on her sense of *charity*—Alice presumes that other people agree with her on the types of things that she takes to be true. This assumption cannot be established by perfect prediction, nor by scientific investigation, but must be assumed as a starting point for translation.

WONDERLAND AND THE REAL WORLD

At the end of her adventures, Alice wakes up and returns to a world more similar to the one that we are familiar with. It would be a mistake, however, to suppose that Alice need leave behind all she has learned about Wonderland, or that her experiences have nothing to teach us. Wonderland presents us with a weird sort of parallel earth where our expectations fail us and we must figure out everything anew. In doing so, we come to see the importance of inductive reasoning for surviving our daily lives, for uncovering the nature of the world around us, and for understanding one another. These are projects that we, like Alice, cannot help but care about.

As readers, we can look to Alice for a model of what it takes to be a good inductive reasoner. Alice, unlike the strange creatures she encounters, learns to control her size, and does not fall prey to the fatalistic belief that she lives only in the dream of another. In contrast to the jurors in the knave's trial, Alice realizes how preposterous the interpretation of the note really is. And she does all of this in the face of repeated challenges to her justification for reasoning as she does. In the end, Alice serves as a good reminder that not *all* reasoning is pointless or silly, and that a good reasoner is much more than an expert in the verbal gymnastics practiced by the residents of both Wonderland and our own world.

1 Lewis Carroll. *Symbolic Logic* (London: MacMillan, 1896), p. 64.

2 Lewis Carroll. "Alice's Adventures in Wonderland". In: *The Annotated Alice*. Ed. by Martin Gardner (Bungay: Penguin Books, 1965), p. 31.

3 David Hume. *An Enquiry Concerning Human Understanding*. Ed. by Eric Steinberg. 2nd ed (Indianapolis: Hackett Publishing Company, 1993), p. 14.

4 Lewis Carroll. "Through the Looking Glass". In: *The Annotated Alice*. Ed. by Martin Gardner (Bungay: Penguin Books, 1965), p. 269.

5Hume, p. 15.

6 Quine discusses this problem in (W.V. Quine. *Word and Object*. Cambridge, MA: M.I.T. Press, 1960) and (W.V. Quine. "Two Dogmas of Empiricism". In: F*rom a Logical Point of View*. Cambridge, MA: Harvard University Press, 1980, pp. 20-47)

7Carroll, "Through the Looking Glass", p. 239.

8Quine, "Two Dogmas of Empiricism", p. 43.

9Quine, *Word and Object,* pp. 19-22.

10Carroll, "Through the Looking Glass", p. 341.

11Quine, *Word and Object*, pp. 52-53.

12Carroll, "Alice's Adventures in Wonderland", p. 103.

13Carroll, "Alice's Adventures in Wonderland", p. 159.