Plot Twists and Surprises:
Why Are Some Things Improbable?

Alfred Hitchcock's movie *Foreign Correspondent* prompted one film reviewer to give this succinct summary: "More twists than a cruller." We could say the same thing about the unfolding plot lines in most of Hitchcock's movies. When you watch one of his films, you might see an ordinary citizen become entangled in international intrigue through a series of improbable events. Or you might come across a surprising revelation that you never saw coming. But one thing is certain. In some form or another, you'll see the improbable, the surprising, the shocking.

But exactly what makes an event improbable? Philosophically, this has become an interesting question in light of the distinctions twentieth-century philosophers have made between different *kinds* of probability. In short, there are different *ways* in which an event can be improbable. And Hitchcock's films are a rich resource of examples of events that are improbable in these various senses of "probability." If we look at the plot developments from a few of his classic movies, we can see how they illustrate the different kinds of probability philosophers have distinguished.

The Science of Frank Fry's Jacket

Much of the modern-day discussion of probability builds on the original work of seventeenth-century philosophers such as Blaise Pascal (1623–1662) and Gottfried Leibniz (1646–1716). Philosophers since that time have not always agreed on how to distinguish the different kinds of probability. Still, by applying what philosophers have said about probability, we can distinguish three general ways in which an event can be improbable.

First, there is what we might call *physical* probability. This is the extent to which a particular event is determined by previous natural events. In past centuries, many philosophers and scientists thought the world worked "deterministically." They imagined the world to be like the internal machinery of a clock. Once a clock is wound up and begins to tick, the cogs of one revolving wheel turn the cogs of another wheel. And the cogs of this second wheel cause the cogs of yet another wheel to turn. And so on. This scheme allows no room for random or chance events; everything that happens is completely determined by prior physical events.

Hitchcock at times seemed to hint at the possibility that our world doesn't operate within such a simple, deterministic mechanism. Consider the bizarre behavior exhibited by the winged creatures in *The Birds*. While it is *possible* that there are prior physical events that cause the birds' behavior, Hitchcock deliberately attempts no such explanation. And so he leaves the door open to the idea that things in our world sometimes occur due to mere freakish, physical chance.

Current scientific research supports the idea that there are some things in our world which are not mechanically determined by prior events. Certain kinds of carbon atoms, scientists tell us, decay at different rates. And there's simply no way to tell when they'll decay. The most we can say about any particular *collection* of carbon atoms is that it has, for instance, a fifty-per-cent chance of decaying within the next five thousand years or so. To cite another example, sub-atomic particles like electrons can move in ways that, from all appearances, are completely random. When scientists fire an electron directly at a target on a screen, the most they can say is

that the electron will hit *somewhere* on the screen—with, for instance, a twenty-percent chance of hitting the specified target.

These kinds of 'indeterminate' events which scientists have documented are all at the micro-level of atoms and sub-atomic particles. It's an open question whether our larger, macro-level world of bowling pins and billiard balls operates deterministically or indeterministically. Our everyday observations of the world *seem* to suggest that things operate deterministically on the macro-level. I've never shot a billiard ball toward a corner pocket only to have it randomly swerve to a side pocket (as much as I'd like to be able to use that as an excuse!). I'm guessing you would say the same. Still, some philosophers of science theorize that the indeterminacy at the micro-level means that the macro-level is also in principle indeterministic—even if it is extremely unlikely that enough micro-particles would ever move randomly in the same direction so as to make a billiard ball suddenly swerve or make a bowling pin unexpectedly spin and topple over.

Because physically improbable events seem only possible at the micro-level, this first kind of probability—physical probability—is of little interest for our larger discussion. In the movie *Sabateur*, Robert Cummings's character Barry Kane grabs the jacket sleeve of the baddie Frank Fry in an attempt to keep Fry from falling from the top of the Statue of Liberty. We then watch the stiches of the shoulder seam of Fry's jacket tear one by one from the weight of the dangling villain. Slowly, agonizingly, the seam finally comes apart altogether and Fry plummets to his death. If, at the macro-level of objects like jackets, the world does in principle work indeterminately, then it will be a matter of probability whether the stitches of the coat hold. Specifically, this will be a matter of *physical* probability. But, again, this is not the sort of probability that is likely to account for the unlikely and surprising events we might witness in the world. So let us turn to the two remaining kinds of probability.

Statistics for Umbrellas?

These last two kinds of probability are given various names in the philosophical literature. Rudolf Carnap (1891–1970), who influentially wrote about the distinction between these two kinds of probability, gave them the names "probability$_1$" and "probability$_2$." Creative, huh? Perhaps we can do a little better! Let's call one *statistical* probability. This refers to the proportion of particular events that are contained within a larger class of events. Suppose in a certain voting district forty-five percent of the people are registered Rebublicans, fifty percent are registered Democrats, and five percent are independent voters. If you were to conduct a phone survey, how likely is it that you would reach a Republican if you dialed a random phone number? You'd of course have a forty-five percent chance of reaching a Republican. And this would be a matter of *statistical* probability. If you happened to reach an independent voter, this would be a statistically improbable event.

In the movie *Foreign Correspondent* a classic scene takes place on a rainy Amsterdam street where a gathered crowd of people are all holding black umbrellas. An assassin shoots a political leader named Van Meer and then disappears into the crowd. Joel McCrea's character, Huntley Haverstock, tries to locate the assassin in the crowd. But there is a sea of identical-looking black umbrellas before him—and only one assassin, who could be hiding under any one of them. Unless Haverstock could somehow search under the majority of the umbrellas, it's unlikely that a simple umbrella-by-umbrella search would yield the assassin. The unlikelihood here is one of *statistical* improbability.

A similar scene takes place in *North by Northwest*, where Cary Grant's character Roger Thornhill puts on a train porter's uniform in order to sneak past the police, who are looking for him. When the police discover that Thornhill has stolen the uniform (well, he actually bribed a porter for it), they begin searching the train station for him. Their "search" amounts to accosting anyone they happen to see in a porter's uniform. But with so many porters at work in the station, it is very unlikely that any single uniform they spot will be the one Thornhill is wearing. And in the end he is able to escape before the police locate him. The unlikelihood of finding Thornhill among the multitude of similarly-dressed porters is again a matter of statistical improbability.

A particularly interesting case of statistical probability occurs in *The Wrong Man*. A villain commits a string of armed robberies. This is particularly bad news for Henry Fonda's character, Manny Balestrero. For the villain looks remarkably similar to Manny, and eyewitnesses to the crimes identify Manny as the one who committed the robberies. Moreover, when the police question Manny and ask him to write down the words used in an original note by the armed robber, they find that Manny has remarkably similar handwriting. Inexplicably, he even makes the same spelling mistake that the robber made in writing the original note. Hitchcock offers a commentary at the beginning of the movie: "This is a true story, every word of it; and yet it contains elements that are stranger than all the fiction that has gone into many of the thrillers that I have made before." The main story line is improbable in the following sense. We ordinarily see many, many people in the course of our lives. And rarely, if ever, do we meet two people (especially two unrelated people) whom we cannot tell apart—and who even have similar handwriting and are prone to the same careless spelling mistake. So, experience tells us that it would be extremely unlikely that Manny would have a "double" in all these respects.

Of course, with a world population in the billions, it is perhaps not that unlikely that *someone* somewhere in the world has a double who could fool multiple eyewitnesses, and who is prone to the same spelling mistakes. And Hitchcock, with his "ordinary man" motif, has a way of reminding us of the unsettling possibility that the next strange coincidence in the world just might throw *our* lives into chaos.

Still, it remains hugely improbable that you, or I, or Manny Balestrero, or any specific individual should have a double somewhere in the world. Perhaps it's not all that unlikely that *a* person somewhere in the world has a double. But it is still wildly improbable that Manny Balestrero would have a double. The improbability here is, once again, a statistical matter.

Show Me the Evidence!

Our third and final kind of probability is what we might call *evidential* probability. This is a matter of how probable a theory is, given a certain collection of evidence. Juries must make probability assessments of this kind. They look at the available evidence and evaluate the competing theories offered by the prosecutor and the defense attorney, respectively. In criminal cases, juries must also determine whether the evidence points "beyond a reasonable doubt" to the prosecutor's theory being true (and thus to the defendant being guilty). In civil cases, jurors need only determine whether the claimant's version of events is more probably true than the defendant's version.

In the previous section we saw that it was a matter of *statistical* improbability that Manny Balestrero from *The Wrong Man* would be mistaken for someone else. But the jury at Manny's trial was concerned with the *evidential* probability that he was guilty. It's true that the jury might have considered the statistical improbability of physical "doubles" as they reviewed all their

pieces of evidence. But it was the evidential probability of Manny's guilt which they were assessing.

What considerations are relevant in assessing evidential probability? This question may draw a mixed response from philosophers, in virtue of an ongoing debate about whether there are *correct* methods of evidence assessment. It seems clear, though, that in our everyday life we assume that there are indeed correct and incorrect ways of assessing evidential probability. Suppose I go with some friends to the movies one evening and I don't bother to lock my car in the parking lot. When asked why I didn't lock the car, I respond by saying that I've already had two cars stolen in the last year. Surely, I explain, the "law of averages" means that it would be extremely unlikely that I would have a *third* car stolen.

In such a case, you would be right to point out to me that I have a thoroughly mistaken understanding of the "law of averages." My statements suggest I think it a statistical matter that car owners who experience thefts are unlikely to experience yet more thefts. Yet, it is obvious that this so-called "law" I cite does not provide the evidential support for my theory (that I won't have my car stolen) that I think it does. My methods for evidence assessment are sadly incorrect. So, there *do* seem to be correct and incorrect methods for determining evidential probability.

What are the correct methods for assessing the evidential support for a theory? Evidence is related to the probable truth of a theory in three ways. So, we must ask three questions about the evidence. It is only by addressing all three that we can correctly evaluate all the ways evidence relates to a theory. Here are the three questions:

(1) Does the evidence fit the theory?
(2) Does the evidence fit *multiple* theories?
(3) Is the theory what we would have expected beforehand?

Jimmy Stewart: Amateur Sleuth or Philosophical Genius?

A concrete example of correct evidence assessment will help us see the relevance of these three questions. Fortunately, we have just the example we need in *Rear Window*. In this Hitchcock film James Stewart's character, L.B. Jeffries (or Jeff), suspects that his neighbor, Lars Thorwald, has murdered his wife. Jeff calls on an old friend, police detective Thomas J. Doyle, to help investigate. Jeff and Doyle then engage in a series of heated discussions about the likelihood that Thorwald is guilty. In their discussions, Jeff shows himself to be quite the amateur sleuth in that he demonstrates an appreciation for each of the questions listed above.

The first question assumes (for the sake of argument) that the theory in question is true. It then asks us whether, on this assumption, we would expect to see the same collection of evidence that we have in fact seen. Suppose the theory is that Thorwald murdered his wife. The first question asks: *If* we assume that Thorwald did murder his wife, would we expect there to be the evidential clues of the same sort that Jeff observed? In short, does the evidence fit the theory?

In his debate with Doyle, Jeff wants to focus on Thorwald's mysterious trips at night in the rain, his wrapping up knives and a saw in newspaper, and his wiping down the bathtub. Doyle, on the other hand, wants to balance the discussion by reminding Jeff of such matters as the superintendent seeing the Thorwalds leave the apartment at 6:00 a.m., Mrs. Thorwald being seen on the train to Merritsville, and a postcard that was mailed from Merritsville back to Thorwald later that day. In sum, Jeff is focusing on evidence we *would* expect to see, on the assumption

that Thorwald is guilty. Conversely, Doyle keeps bringing up the evidence we would *not* expect to see, on the assumption this theory is true.

At one point, Jeff and Doyle debate the same piece of evidence, but they give it different interpretations. Noting that Thorwald's shades are open and that he's sitting in his apartment, Doyle says, "It's too stupid and obvious a way to murder—in full view of fifty windows—and then sit over there, smoking a cigar, waiting for the police to pick him up….That salesman wouldn't just knock off his wife after dinner, toss her in a trunk, and put her in storage." Doyle is arguing that this is not the kind of evidence we would expect to observe, if Thorwald really had murdered his wife. But Jeff insists that it *is* what we would expect, explaining at one point: "That's where he's being clever. Acting nonchalant." Their debate is precisely on the first question listed above: whether the evidence fits the theory.

The second question asks us whether the evidence fits with multiple theories. Put another way, it asks us how likely it is that we would observe the evidence of the case, *whether or not* the theory is true. To see the importance of this question, think about where we stand so far. In answering the first question, we might conclude that Thorwald is probably guilty because the evidence of the case is very much what we would expect to see if Thorwald did in fact murder his wife. But this consideration loses its force if we would expect to see this same evidence, *even if the murder didn't occur*.

Jeff rightly argues that, on the theory that Thorwald murdered and disposed of his wife, we would expect to find him doing such things as cleaning up a saw and washing down a bathtub. But suppose for a minute that everyone in that community cut up wood for their stoves each morning. In such a case, *most* people might clean their saws every day. So, seeing Thorwald cleaning his saw wouldn't constitute telling evidence that he committed a murder. Or suppose that in their community everyone washed down their bathtubs every day to prevent the spread of a bacterial epidemic. We would expect to see Thorwald washing down his tub *whether or not* he hacked up his wife. Thus, the evidential support for our theory (that Thorwald murdered his wife) stemming from Thorwald's behavior would be undermined to the extent that we would expect to see this behavior *even if* our hypothesis is false. The second question—Does the evidence fit multiple theories?—factors in this point.

Jeff and Doyle come to very different conclusions about how likely it is that the evidence of the case would exist once we exclude the theory that Thorwald is guilty of murder. Jeff insists that there is no reasonable, innocent explanation for Thorwald's observed behavior. He asks, "Why would a man leave his apartment three times, on a rainy night, with a suitcase? And come back three times?" Jeff is emphatic that this is *not* what we would ever expect to observe, if no murder had occurred. He points out that Thorwald's wife is "an invalid who needs constant care." Surely, he reasons, we would expect to see Thorwald taking greater steps to care for her—if we assume for the sake of argument that she is still alive.

Doyle, on the other hand, insists that it *isn't* far-fetched that we would see the events involving Thorwald's behavior, if we consider only alternative theories that do not involve Thorwald being a murderer. Jeff objects to this idea, saying: "You mean you can explain everything that went on over there—and is still going on?" Doyle responds, "No, And neither can you. That's a secret and private world you're looking into out there. People do a lot of things in private that they couldn't explain in public." Doyle's point is that, even though there is an air of "mystery" to Thorwald's behavior, it's still the kind of innocent activity we can expect to see in our varied, complex world.

So, Jeff and Doyle give different answers to our second question related to correct evidence assessment. Jeff thinks it very improbable that the evidence of the case would exist, if we assume for the sake of argument that Thorwald didn't murder his wife. Doyle disagrees. But they both recognize the importance of our second question: whether the evidence fits with multiple theories.

The third question concerns the extent to which our prior, background knowledge leads us to think that a theory is likely to be true. When juries consider whether a defendant is innocent or guilty, they do more than consider the specific evidence of the case (such as fingerprints on the murder weapon, and eyewitnesses to the suspect's whereabouts). Jurors also consider their background knowledge of the defendant. Is the defendant a habitual criminal, who has shown he has no regard for human life? Or is the defendant a model citizen, who has consistently demonstrated strong moral character and an aversion to violence? Our background knowledge of the defendant will affect how much specific evidence we need in order to think him probably guilty of the crime with which he's been charged.

Suppose I told you that I've heard various people report that Hitchcock, before he died, was planning to direct a lavish musical remake of *Gone with the Wind*. Surely you'd be skeptical. Yes, Hitchcock did direct the comedy *Mr. and Mrs. Smith* and the farcical *The Trouble with Harry*. But at no time did he ever show any real interest in directing musicals or in plot lines that resemble lavish soap operas.[1] Surely, you'd think to yourself, this must be some crazy Internet rumor. Your background knowledge of Hitchcock is so at odds with my testimony that this testimony is not nearly enough evidence to convince you that my theory is likely true.

The third question of evidence assessment factors in this point and asks us to consider how likely a theory is true, given only our background knowledge. Doyle has no strong views here about Thorwald. His investigation of Thorwald yields only the following: "He has a six-month lease…Quiet. Drinks, but not to drunkenness. Pays his bill promptly…Keeps to himself, and none of the neighbors got close to him, or his wife." As a detective, Doyle is not naïve enough to think that quiet men never murder their wives. But most men don't go around seriously considering whether to murder their wives. And since Doyle sees nothing in Thorwald's background that makes him think Thorwald is different from most men, he doesn't think his background knowledge makes it at all likely that Thorwald is a murderer.

Jeff, on the other hand, thinks there's more to be said about Thorwald's character and his pre-existing relationship with his wife. He says at one point, "I've seen things through that window! Bickering, family fights…" When Thorwald seeks advice from someone on the phone, Jeff reasons from his background knowledge of the couple that it couldn't be the absent Mrs. Thorwald: "I never saw him ask for *her* advice before. But she volunteered plenty!" Jeff is saying here that the background knowledge available about Thorwald *does* offer some reason to think he might have a murderous character. Perhaps Jess wouldn't go so far as to say this background evidence gives a *strong* indication that Thorwald has the kind of vicious character necessary for cold-blooded murder. But considering only background knowledge, he at least thinks the theory of Thorwald being a murderer is not so unlikely as Doyle seems to think it is.

So we've seen that Jeff and Doyle give different answers to each of our three questions about how evidence relates to the probable truth of a theory. It is thus understandable that they reach such contrasting conclusions about the likelihood of Thorwald's guilt. Moreover, there remains

---

[1] It's true that, early in his career when he was without a picture to direct, Hitchcock agreed to do *Waltzes from Vienna* so that he could keep working. He also announced to the cat on set that such a movie went against his instincts.

the factor—which we've not discussed—that Jeff and Doyle might have particular biases that affect how they answer the three questions. Jeff, for example, by focusing on Thorwald to the point of near obsession, is possibly running away from an issue he needs to confront. Jeff fears his life would be cramped if he were to marry his girlfriend Lisa. He's terrified of the commitment and its implications. Down deep he wants, perhaps, to find confirmation that marriage really is so confining that it drives a man mad.

If this analysis is at least partially correct, then Jeff's reasonable objectivity in evaluating evidence may be compromised. He may really *want* Thorwald to be guilty. And experience shows that people often believe things more readily when they want them to be true.

Still, while these are interesting issues to explore, they involve the ways we might be biased in *how* we answer our three questions. Our main concern, though, has been to identify *which* questions need to be asked in assessing evidential probability. And on this point Jeff and Doyle agree. The fact that Jeff and Doyle debate each of our three questions shows that, though they disagree about how to interpret the evidence, at least their *method* for assessing evidence is correct.

What surprises us in Hitchcock movies is often the occurrence of an event we thought was evidentially improbable. And Hitchcock's characters are themselves frequently surprised by evidentially improbable events. We can think of Patricia Martin's character Priscilla Lane in *Saboteur* being told that Barry Kane, who is on the run from the police and who had virtually kidnapped her, really is innocent. Or Jimmy Stewart's character Scottie in *Vertigo*, who is stunned when the woman whose death he is mourning turns out to be alive. And of course Roger Thornhill in *North by Northwest*, who is surprised and angered when he learns that Eve Kendall is in league with Philip Vandamm—and then subsequently surprised and delighted when she turns out to be working undercover for the CIA. This list could go on and on. And the surprise in each case arises when a character discovers something that all his or her previous evidence suggested was not true. In short, the surprise stems from the occurrence of evidentially improbable events.

One memorable instance of this kind of improbability occurs in *The Thirty-Nine Steps*. Robert Donat's character Richard Hannay is being chased by members of a spy ring, whose leader, Hannay has learned, is a clever and ruthless man with a missing right little finger. A woman, just before she's killed by the spy ring, tells Hannay that a man in Scotland may hold the key to stopping the spies. Hannay tracks down the man, Professor Jordan, at his Scottish estate. Relieved at having made it to Professor Jordan safely, Hannay is able to relax completely after the Professor shields him from police discovery and then offers him a drink and a cigarette. Finally at ease, Jordan opens up to the Professor and tells him all that has happened and that the leader of the spy ring has a missing little finger. The Professor asks him which hand has the missing finger, and Hannay says it's his left one. Chillingly, Professor Jordan raises his right hand, revealing a missing little finger, and asks, "Sure it wasn't…*this* one? The shock felt by Hannay and by the audience once again stems from the fact that all our evidence seemed to suggest that Hannay, having reached Professor Jordan, is now safe at last. But as it turns out, just the opposite is the case.

Lessons to Take Away

Keeping the different kinds of probabilities straight can help avoid errors that even professional philosophers have sometimes made! One example involves a critique of arguments from design,

which purport to show the probable existence of God. Noting the number of highly unlikely events that needed to occur for our evolving universe to sustain life, some believers have argued that a creator God is a more probable explanation for the existence of our universe than are purely naturalistic explanations. However, sometimes philosophers have objected to this line of argument on the grounds that there has only been one Big Bang. And because we therefore don't know how many times Big Bangs naturally produce life, we can't say that life in our universe *is* unlikely. But this is a confused objection. The original line of argument is one that compares the *evidential* probability of life in a purely naturalistic world versus a world where God exists. It looks at the evidence of life in our world, and then compares competing explanatory theories. But the critic's objection is that this line of argument doesn't provide *statistical* data of probability. And this is no good objection.

In criminal trials we would never require a jury to have statistics of how many times the defendant has turned out to be truly guilty when previously put on trial for similar crimes. And this is because juries assess evidential probabilities, not statistical probabilities. Similarly, it is improper to insist of the believer, whose argument is based on evidential probability, that he have statistics about how many times similarly evolving universes have and haven't produced life. So, one lesson to keep in mind from our discussion is that we can make errors of reasoning when we do not distinguish the different kinds of probabilities.

A second lesson worth keeping in mind is that a proper assessment of evidential probability requires us to ask three separate questions. Otherwise, we will fail to appreciate all the ways in which evidence relates to a theory. In our discussion of *Rear Window*, we saw that Jeff, without realizing it, engaged in reflection on all three questions. These questions provide more than guidelines as to how evidence *can* be assessed. They tell us how evidence *should* be assessed. And this is important for anyone sitting at home with binoculars, with suspicious-looking neighbors, and with aspirations of becoming an amateur sleuth like Jeff.