This thesis by Willa Maxine Tracy is accepted in its present form by the Department of Philosophy as satisfying the thesis requirement for the degree of Master of Arts.

Date: 5/21/19

signed: Joshua Schechter, Advisor
Table of Contents:

introduction......................................................................................................................4
first role of chance: the Principal Principle.................................................................5
second role of chance: chance & explanation.............................................................6
anomaly & coincidence.................................................................................................8
surprise..........................................................................................................................10
chance & randomness.................................................................................................12
the Shooting Room....................................................................................................14
synthesis.......................................................................................................................16
the problem of counterinduction...............................................................................17
induction & metacognition.........................................................................................21
references....................................................................................................................28
The concept of objective chance plays two indispensable roles in our predictive and explanatory practices. The Principal Principle implements chance as a guide to rational prediction; we additionally rely on chance to explain observations of stable long-run frequencies. These two uses of chance, one forward-looking and one backward-looking, set the stage for a particularly thorny instantiation of the problem of induction. I say “particularly thorny,” because our reasoning about chance hinges entirely on our observations of patterns, observations which are susceptible to distortion by human pattern-seeking tendencies. The two roles of chance facilitate our anticipation and explanation of scientific observations, thus reducing feelings of surprise associated with unanticipated or unexplainable occurrences. We nevertheless feel surprised when chance events do not cohere with our evolutionary inclination to latch onto patterns. Surprising observations should not steer us toward alternative hypotheses of objective chance, because the pattern-seeking aspect of human psychology gives rise to irrational feelings of surprise, in reaction to observations that are in fact consistent with the laws of objective probability. I therefore advocate resisting our tendency to let surprise govern the role chance plays in our predictive and explanatory practices.

In my argument, I will first outline the predictive and explanatory roles of objective chance, which we implement to reduce feelings of surprise associated with unanticipated or unexplainable observations. I will evaluate our experience of surprise, when occurrences seem to conflict with our beliefs about objective chances, and I will present two arguments against a surprise-governed grasp on objective chance. First, the conceptual relationship between chance and randomness does not aptly cohere with our human pattern-seeking tendencies. Our predictive and explanatory practices are readily misguided, when we fail to recognize instances of anomaly and coincidence. Furthermore, the Shooting Room paradox involves divergent beliefs about objective chance, dependent on subjective vantage point. I will implement this
paradox to support my argument that feelings of surprise ought not govern beliefs about objective chance.

first role of chance: the Principal Principle

Among the various classifications and subcategories of chance that may be defined in various contexts, our understanding of chance separates into two basic categories. Rudolf Carnap defines probability as an epistemic notion, referring to evidential probability or degree of belief.¹ In contrast with subjective credence, probability is a non-epistemic objective concept, what we commonly call chance. In what follows, chance will always be used in the probability sense. It is controversial whether chance exhibits temporal asymmetry, where past events can no longer be considered chancy. I side with David Lewis on this matter, as I agree that there can be no present chance of a past event occurring differently. Chance is a future-oriented concept, pertaining to the objective probability of a given event’s occurrence. In contrast, explanation is primarily a past-oriented practice, describing observed events in causal terminology.

David Lewis’s Principal Principle concerns the relationship between chance and credence.² For example, if I were to ask you to describe your credence that a fair coin flip would land heads, you should report a credence level of 0.5. This credence level is conditional on your belief that the coin is fair, where fair means something like, “the chance is 0.5.” A rational agent should defer to chance in this way when forming an opinion about some future event, adopting the corresponding chance as a conditional degree of belief. The Principal Principle yields an unconditional degree of belief, by summing over alternative hypotheses about the relevant chances. I will provide an example of such a summation:

Suppose I have two coins in my pocket. One is a fair coin, with a 0.5 chance of landing heads. The other is a weighted coin, with a 0.9 chance of landing heads. Were I to pull a coin

out of my pocket and flip it, your credence that the coin would land heads should correspond to
the chance of that outcome, weighted according to these two hypotheses about the chances.
There is a 0.5 chance that I will pick the fair coin, which has a 0.5 chance of landing heads.
There is a 0.5 chance that I will pick the weighted coin, which has a 0.9 chance of landing heads.
Summing over these alternative hypotheses, your credence that the coin will land heads should
be \((0.5 \times 0.5) + (0.5 \times 0.9) = 0.7\).

Lewis raises some doubts as to whether the Principal Principle qualifies as an analysis of
chance.\(^3\) The principle advocates conforming subjective credence to the objective chances,
which would require some sort of independent access to knowledge of what the chances are.
The question remains how we might form accurate true beliefs about objective chance, in order
to implement the Principal Principle to make predictions.

*second role of chance: chance & explanation*

I previously alluded to the temporal asymmetry of chance, where a past event can no
longer be considered chancy. The prior probability of an event’s occurrence seems to play a
necessary explanatory role in our understanding of why that particular outcome occurred. In this
context, we do not consider the present probability, because a past event has no chance of having
occurred differently. Instead, the prior probability of an event’s occurrence stands in an
explanatory relation to the factual outcome. Nina Emery proposes the concept of nomological
probability, defined by its explanatory role with respect to stable long-run frequencies.\(^4\)
Nomological probabilities are objective, determined by physical features of the world, and
independent of the epistemic position we might occupy. Nomological probabilities are not actual
frequencies; they are magnitudes playing a certain role in our explanatory practice. I will give an
example of a stable long-run frequency, explained in terms of nomological probability:

---

Suppose I have a hole in my pocket and I lose my weighted coin from the previous example. You find the coin, and you notice that it seems unusual in some way. You rightly suspect that it is weighted, so you begin flipping the coin and recording the outcomes of either heads or tails. After one hundred flips, you observe the surprising result of ninety-two heads and only eight tails. This stable long-run frequency calls for some explanation. You had suspected that the coin was weighted, and your recorded observations seem to suggest that the objective chance of the coin landing heads is approximately 0.9. You continue flipping the coin and the frequency of its landing heads persists. In this instance, this robust pattern of events (i.e. the weighted coin landing heads with the stable long-run frequency you observed) is explained by the corresponding nomological probability of 0.9 for a heads outcome.

The explanatory power of nomological probability depends on a departure from actual frequency analysis, which equates objective probability with frequency. Because nothing can explain itself, nomological probability corresponding to chance must stand as an independent concept from the frequencies that characterize robust patterns of events. The necessary independence of these concepts is evidenced by the explanatory gap that emerges when the outcome of a chance process strikes us as surprising in some way, indicating a call for explanation beyond that provided by actual frequency analysis.

As Emery argues, nomological probabilities play a necessary explanatory role, despite the accompanying metaphysical scruples. These probabilities explain observed patterns of events, which cannot be explained by any other means. The standards of naturalism uphold scientific methodology in pursuit of successful inquiry, where the norms and constraints guiding scientific theory oversee metaphysical theory in turn. It would therefore be costly to our naturalistic principles not to explain observations of robust patterns of events, which remain consistent under a variety of temporal, spatial, and counterfactual conditions. I uphold Emery’s naturalistic position that this cost outweighs that of novel metaphysical entities posited by an explanation, i.e. nomological probabilities. Necessary explanatory entities should therefore have priority over our accepted metaphysical principles. Nomological probabilities are sui generis entities, which cannot be identified with, or analyzed in terms of, more familiar classes of
metaphysical entities. It is difficult to pin down exactly what nomological probabilities are; they can only be grasped through their necessary explanatory role.

The predictive and explanatory functions that I have outlined seem to mutually constitute the concept of chance. Our observations of stable long-run frequencies call for an explanation in terms of objective chance, which in turn informs our credence level that those patterns will continue. Our predictive and explanatory practices are respectively forward-looking and (primarily) backward-looking, seeming to temporally complement one another.

anomaly & coincidence

Feelings of dissatisfaction with a chance explanation can arise in two types of scenarios, which I will call anomaly and coincidence. An anomaly is a result that contradicts some expected pattern of observations. For example, if I had a weighted coin, whose chance of landing heads was close to 100%, the coin landing tails would be an anomalous result. An even more striking anomaly, would be if the coin started levitating. Anomalous states of affairs are points of deviation from a stable long-run regularity, which seem to call for explanation. Possible explanations for the levitating coin might include the coin being raised by a magnet, or an undetectable string, or the coin being a hologram. The pressure we feel to settle on one of these explanations comes from our firm belief that the laws of physics guarantee certain regularities of observation, which conflict with the anomalous levitating coin.

Coincidence, in contrast, is the emergence of a pattern where one was not expected. Flipping a fair coin one hundred times with a sequence of ninety-two heads, would therefore be a coincidence. Both anomaly and coincidence are typically accompanied by a feeling of surprise. This feeling, when strong enough, motivates our search for a new, more satisfying explanation. We might suspect that the sequence of ninety-two heads is due to the coin being weighted, or some biased flipping technique. Although it is possible for a fair coin to land in this sequence, the low probability of this result motivates the search for some alternative hypothesis that favors this seemingly coincidental result.
Martin Smith, however, argues that a sequence of coin flips landing heads ninety-two times in a row is not at all surprising.\textsuperscript{5} He reasons from the premise that one should never feel surprised at an individual result of a chance process: it would be irrational to react to the outcome of a single coin flip or dice roll with any degree of surprise. These chance processes entail that each possible outcome has an equal probability of occurring, therefore any resulting outcome bears no relational status regarding its likelihood, compared with other potential outcomes. Smith argues that the inherent lack of surprise associated with any single chance event might lead to problematic implications under the conjunction principle. This principle states that if two independently occurring events are separately unsurprising, the conjunction of these two events must also be unsurprising.

I find this conclusion is unacceptable, because a streak of ninety-two heads constitutes evidence suggesting some causal mechanism at work, beyond the allegedly unbiased chance process of coin-flipping. I would therefore propose revising the conjunction principle to require not only the antecedent belief that separately unsurprising individual events occurred independently of each other, but also the continued certainty that a conjunction suggests no relationship of causal dependence among the conjoined events. A sequence of ninety-two heads in a row fails to satisfy the second requirement of my revised principle, because this result undermines the certainty that these events occurred independently. The conjunction of ninety-two consecutive coin flips landing heads suggests a relationship of dependence among the results of each individual coin flip.

Anomaly and coincidence indicate explanatory urgency. The subjective experience of surprise at an outcome provides some motivation to develop an alternative hypothesis, involving further causal mechanisms beyond the explanatory framework proposed by the null hypothesis. While I believe that Smith’s argument is wrong, and a coin landing heads ninety-two times in a row would be rationally surprising, our feelings of surprise do not reliably indicate a true demand for explanation. The significance of a given surprising event is bound up in questions of what corresponding beliefs we ought to hold, where rational surprise is a guide to rational belief. It can be challenging to distinguish between events that are in fact rationally surprising, and

\textsuperscript{5}Smith, Martin (2017) Why throwing 92 heads in a row is not surprising. _Philosophers’ Imprint_ 17.
coincidental events that merely appear to call for explanation. The felt need to revise a chance explanation (either by recalculating the chances, or positing some additional causal forces) is not necessarily justified. I therefore advocate resisting our tendency to let surprise govern our explanatory practices. Before detailing the problem with surprise-governed explanation, I will first say a bit more about our explanatory methods, and the relationship between surprise and coincidence.

surprise

A surprising event that is insufficiently striking to justify believing an alternative theory creates the illusion of explanatory urgency. A coincidence induces feelings of surprise, which are informed by our irrational biases toward patterns of past observations. Human beings are evolutionarily pattern-oriented; we make inductive inferences based on our habit of adhering to the patterns of past occurrences, despite the lack of rational justification for believing those patterns will continue. The ability to form predictions about the future is evolutionarily advantageous, but our inductive practice cultivates an inclination toward uncovering the existence of patterns, even when no pattern is being causally generated. Failure to dismiss certain surprises as mere coincidences can lead us to overestimate the plausibility of novel causal forces, and form false beliefs about laws governing our observable universe.

Our perception of a coincidence as significant does not directly correspond with probability, as illustrated by our inconsistent feelings of surprise associated with events of equal probability. A fair coin landing heads ninety-two times in a row is a suspicious coincidence, bringing to mind alternative hypotheses regarding potential bias of the coin. A recognizably random sequence of outcomes for ninety-two consecutive coin flips would not share those causal implications, even though that unique random sequence had the same expected probability as a sequence of ninety-two heads. A coincidence suggests the presence of a hidden causal structure, in contexts where the current understanding suggests that no such structure should exist.
Detecting a coincidence therefore does not only involve recognizing an unusual pattern, but doing so despite the presence of some observations that do not express that pattern.

A coincidental event inspires conflicting feelings of surprise. An event that seems surprising in the context of the currently accepted causal framework might be construed as evidence supporting an alternative hypothesis. The truth of this alternative hypothesis would itself be surprising, given its low a priori probability. I have chosen to focus my argument on coincidence, rather than anomaly, because a coincidence more explicitly gestures toward some particular alternative hypothesis, which would explain an unexpected pattern of observations.

The support that a coincidence provides for some alternative hypothesis might be insufficient to override our a priori beliefs about the implausibility of that hypothesis. In these instances, the continued belief in a currently favored theory requires an explanation of the surprising event that involves some chance process, which cannot be described in causal terminology. A coincidence becomes more suspicious as it renders an alternative theory more plausible, and can eventually serve as evidence justifying that alternative theory.

Thomas L. Griffiths and Joshua B. Tenenbaum elaborate on Paul Horwich’s theory of coincidences, in the context of a Bayesian framework for causal induction. A coincidence is an event that provides support for an alternative to a currently favored causal theory, but not necessarily enough support to accept that alternative in light of its lower probability. There exists a paradox for theories about human reasoning, where coincidences are both a source of important scientific discoveries, as well as widespread false beliefs.

Coincidences arise when there is a conflict between the evidence an event provides for a theory, and our prior beliefs about the plausibility of that theory. This produces a likelihood ratio that is insufficient to overcome the prior odds against that event occurring, resulting in middling posterior odds. This relationship between the likelihood ratio and prior odds helps distinguish between mere coincidences and suspicious coincidences. The former characterizes surprising events that are ultimately believed to be the work of chance, and the latter refers to events that begin to render an alternative theory more plausible.

---

Events transition from suspicious coincidence to evidence, as the posterior odds increase. Assessing the strength of a coincidence is equivalent to assessing evidence for a causal relationship. Whether an event is a coincidence or evidence for an alternative theory comes down to whether it ultimately justifies believing that theory. We make such assessments poorly, because we are overly sensitive to patterns. As I will discuss in the next section, the subtle properties of product randomness lead our pattern-seeking tendencies astray, disrupting our explanatory and predictive practices. Inaccurate human reasoning about objective chance explains the potential for coincidental events to inspire feelings of surprise, that cannot be rationally defended.

*chance & randomness*

As I have asserted, our reasoning about chance is a particularly thorny instantiation of the problem of induction: we have an evolutionary inclination to form tenuous beliefs about objective chance, based on the patterns we observe. Other instances of the problem of induction, such as our inductive inference that the sun will rise tomorrow, postulate causal mechanisms in the context of scientific theories explaining the continuation of previously observed patterns. Reasoning about chance, in contrast, hinges entirely on our attention to patterns. While we also have background beliefs about causal mechanisms that inform our beliefs about the chances (such as beliefs about the physical properties of a fairly weighted coin), accurate predictive and explanatory application of these beliefs depends on our ability to accurately interpret patterns of observation.

For example, let us again consider your discovery of the weighted coin, where you observe its stable long-run frequency of landing heads about ninety percent of the time. You explain this robust pattern of events in terms of nomological probability. Now looking forward, when the coin is flipped, you ought to have a credence level of 0.9 that the coin will land heads, in accordance with the Principal Principle.
It seems as if, at some point, our understanding of objective chance comes out of actual frequency analysis. While nomological probability should stand independently of actual frequencies, our practical ability to discern probability relies on the frequencies we observe. Subjective credences follow from our patterns of observations, which in turn inform our credence that those patterns will continue. Our desire to predict and explain patterns of observation is frequently misguided by our disproportionate attention to patterns that emerge within random products.

There is a serious problem complicating our use of chance to explain and predict patterns of events. This problem involves our human inability to accurately recognize or generate randomness. A random sequence generated by a chance process exhibits a property that I call “second-order disorder,” where the sequence contains brief runs of apparent patterns, interspersed among more recognizably random subsequences. The human tendency to overlook second-order disorder of random sequences can be evidenced by a simple psychological experiment. If one group of subjects is asked to record the outcomes of one hundred fair coin flips, while another group is asked to invent a sequence of outcomes for one hundred fair coin flips, the results for each group should be easily distinguishable. A subject attempting to generate a random sequence tends to alternate between heads and tails too quickly, but in actuality, a lengthy series of fair coin flips will typically include long streaks of the same result occurring consecutively, a counterintuitive characteristic of randomly generated sequences.

Our disproportionate attention to patterns occurring in the midst of randomness prompts us to form beliefs that hidden causal mechanisms are at work, in instances where anomalous or coincidental events create the illusion of causal dependency. Because we cannot reliably distinguish between explanatorily urgent patterns and mere coincidences, we ought not take our intuitions too seriously. While an explanation that posits some additional causal framework might feel more satisfying, sometimes we must accept that a set of observations occurred by mere chance. Null hypotheses about objective chances are robustly supported by long-run patterns of observable evidence. We should therefore have more inertia to rest with these hypotheses, and override our initial emotional response to surprising observations.
Here, I will present the second problem with a surprise-governed grasp on objective chance. In addition to instances of anomaly and coincidence, we experience the feeling of surprise when our evaluations of objective chance vary drastically, according to our particular epistemic vantage point in a given scenario. Varying degrees of surprise experienced by different subjects ought not influence their beliefs about objective chance. The Shooting Room paradox involves divergent subjective credences about the occurrence of a singular event. I will not present a solution to the paradox, rather I will use it to illustrate disparate subjective experiences of surprise, which should not govern the explanatory role of chance:

You are summoned to a room called the Shooting Room. The Shooting Room is not as foreboding as you had anticipated, when it is revealed that you will only be shot if the rolling of two fair dice lands double sixes. You relax a little, knowing the probability of your death is only 1/36. Using your investigation skills, you deduce that you are not the only subject of The Shooting Room – you are part of a game. In this game, a single subject is called to the room, followed by 9 subjects, then 90, then 900, and so on with each group of subjects in the room at one time representing 90% of the total number of people who have played the game (with the exception of the first round, where the sole subject represents 100% of the game’s players). You do not know what group you are in – the room is dark; but you know that is irrelevant, because your chance of dying can only be 1/36. The game will go on until double sixes are rolled. The population of available players is countably infinite, so once you complete your round, you will not be called into the Shooting Room again. You play your round at the Shooting Room.

In the town of countably infinite people, your mother is reading the newspaper. She reads that the game has ended, and there is a published list of all the players. Your name is on the list because you played the game. Your mother is devastated, because she knows 90% of the players on that list are dead. What information does your mother have that you do not? And vice versa? We all know mothers worry, but this outside the standard deviation. How could a game

---

where each player has a 35/36 chance of survival result in 90% of the players dying? We have reached the paradox.

Let’s take an inventory of the information you and your mother both have: you both know that you participated in the game, and neither of you know what round you were in (although that should not matter). Neither of you know how the dice came up in your round. Your mother, however, knows the game has ended, while you lack this information when you are playing your round. You know the game will eventually end, because there is a probability of 1 that double sixes will eventually be rolled. This happens in the fatal round, where 90% of those who have ever entered the room will be executed. While the game is still in play, there is a possibility that it will never end, but that probability converges at zero. Therefore, your mother’s knowledge that the game has ended cannot account for the paradoxically disparate calculated probabilities of your death.

Imagine that you have played the Shooting Room game and survived. This should not be surprising to you, because you knew that you had a 35/36 chance of survival. Your mother, however, is very surprised to see that you are alive, after learning that you played this game with a 90% fatality rate. Her surprise at seeing you alive should not undermine your belief that you had only a 1/36 chance of dying in the game. You understand the outcome of your survival in terms of the high nomological probability of that result. Because our evaluations of objective chance can vary drastically according to our particular epistemic vantage point in a given scenario, feelings of surprise can arise when considering the chances from a different position in the scenario. These disparate feelings of surprise ought not govern our assessment of the chances; your mother was right to worry about you, and you were right to remain fairly calm.

This paradox involves a peculiar discrepancy of subjective grasp on objective chance, where divergent experiences of surprise should not play a role in explaining your survival in the Shooting Room game.
I have argued for two indispensable roles of objective chance, one forward-looking and one backward-looking, which set the stage for a particularly thorny instantiation of the problem of induction. We try to anticipate and explain the outcomes of chance processes in accordance to sequences of events we observe, a practice that is susceptible to distortion by human pattern-seeking tendencies. This aspect of our psychology can lead to irrational feelings of surprise, in reaction to observations that seem not to cohere with our beliefs about objective chance. We should therefore resist the tendency to let surprise govern our theories of objective chance.

Even if we were able to flawlessly implement the Principal Principle and explanatory probability to accurately predict and explain patterns of observations, events still might surprise us. We could rationally form false beliefs that the world would later demonstrate to be false. Our predictive and explanatory practices do not quite grasp the independent concept of objective chance, therefore our theories of chance depend somewhat on actual frequency analysis. While nomological probability should stand independently of actual frequencies, our practical ability to discern probability relies on the frequencies we observe. Our desire to predict and explain patterns of observation can be misguided by our disproportionate attention to patterns that emerge within random products. This can lead us to irrational feelings of surprise, in response to observations that are in fact anomalous or coincidental.

Additionally, the feelings of surprise involved in the Shooting Room scenario should not govern explanation, because the two subjects, you and your mother, seem to have access to the same information. The two roles of chance I have outlined are used to predict and explain patterns of observation, with the aim of reducing feelings of surprise. However, when confronted with surprising observations, the feeling of surprise alone should not motivate a revised theory of chance. Rather, a better understanding of the conceptual relationship between chance and randomness might provide us with guidelines on when to dismiss feelings of surprise in cases of anomaly and coincidence. Furthermore, we should adhere to our own epistemic vantage point when developing a theory of objective chance, and not allow feelings of surprise to
sway us toward a theory associated with another vantage point, as in the case of the Shooting Room. Surprise should not motivate the search for some alternative explanation of the chances. When choosing among multiple conflicting theories of chance, our feelings of surprise should not govern our decision making. The question remains open as to how we should most reliably form our beliefs about objective chances.

**the problem of counterinduction**

I have argued that we ought not let feelings of surprise govern our predictive and explanatory grasp on objective chance. We cannot reliably distinguish between our irrational feelings of surprise at chance outcomes, and rationally surprising events that call for a revised explanation. I have advocated more inertia to rest with null hypotheses about the chances, which are more robustly supported by long-run patterns of observable evidence than a potential alternative explanation. In the remainder of the paper, I will analyze our predictive and explanatory grasp on objective chance in relation to the problem of induction. I will briefly sketch the problem of epistemic circularity, where rule-circular justification would also support the seemingly irrational practice of counterinduction. I will argue that we in fact reason counterinductively when we attempt to anticipate a surprise, which I will illustrate using the surprise exam paradox. I will tie my discussion of counterinduction to my beliefs about the temporal asymmetry of chance.

I have asserted that our reasoning about chance is a particularly thorny instantiation of the problem of induction. This is the problem that keeps us up at night, until the sun rises in the morning. Our inductive inference that the sun will rise tomorrow postulates causal mechanisms in the context of scientific theories explaining the continuation of previously observed patterns. Reasoning about chance, in contrast, hinges entirely on our attention to patterns, where we have an evolutionary inclination to form tenuous beliefs about objective chance, based on the patterns we observe. While we also have background beliefs about causal mechanisms that inform our beliefs about the chances (such as beliefs about the physical properties of a fairly weighted coin),
accurate predictive and explanatory application of these beliefs depends on our ability to accurately interpret patterns of observation.

I have argued for two indispensable roles of chance in our predictive and explanatory practices. In simple terms, explanations are stories we tell ourselves about the world, in order to better understand it. When an explanatory narrative is not forthcoming, we construct one for ourselves. We all rest assured in our belief that the sun will rise tomorrow, despite our lacking any rational justification for inductive reasoning. The ancient Egyptians understood the sun god to manifest himself in multiple forms, chiefly those of Re during the day, the ram-headed Atum in the evening, and the scarab beetle Khepri at dawn. During the sun’s daily journey across the sky he grew old and weary, then was rejuvenated overnight and reborn, young and strong at sunrise. I see this narrative as an illustration of our human inclination to tell ourselves stories that make sense of the world.

Chance explanations are stories we tell ourselves to anticipate and understand the outcomes of chance processes, which ought not be governed by our irrational feelings of surprise. These feelings, although irrational, are very human. Narratives about chance naturally bring to mind notions of luck and superstition; we let our imagination shape our perception of chance events, in order to cope with irrational feelings of surprise. My superstitious belief that I was bound to win the lottery because of the lucky numbers I chose helps me understand the anomalous result of my winning. Though I was unlikely to win, my ticket had the same chance of winning as anyone else’s. The chance explanation that my ticket won because it had a non-zero probability of being chosen feels explanatorily dissatisfying, just as the explanation of a fair coin landing heads ninety-two consecutive times by mere chance is dissatisfying. The tendency for chance explanations to inspire irrational feelings of surprise or dissatisfaction goes hand in hand with irrational beliefs about luck and superstition that accompany our inductive inferences about chance.

The problem of induction is a problem of epistemic circularity. Inductive inferences from premises establishing regularity in observed instances infer that regularity will hold for unobserved instances. These inferences are reliable only if the world works a certain way, in accordance with Hume’s principle of induction. While the inductivist will maintain that the
principle of induction is supported by harmless rule circularity, the counterinduction parody undermines this justification. Formulated by Wesley C. Salmon in 1957, the proposition that counterinduction is rational can be deduced from prior observations and inference by counterinduction. This parody demonstrates that a rule’s ability to support itself is not sufficient for justification, which undermines the rule circular justification for induction. The success of the parody depends on presupposing the belief that counterinduction is not rational, therefore any argument supporting it must be fallacious. I will argue against this presupposition, on the grounds that we rationally employ counterinductive reasoning when we try to anticipate a surprise. My argument follows from the surprise exam paradox:

Suppose your math teacher tells your class that there will be a surprise exam next week. On Sunday night, you are thinking about when the exam might be. You rule out Friday as a potential day, because it would not be surprising to have the exam on Friday. If you and your classmates were to go to school Monday-Thursday without having the exam, you would be expecting it on Friday, so it would not be a surprise. After you have ruled out Friday as a potential day, the same reasoning applies to Thursday: with no exam Monday-Wednesday, and with Friday already ruled out, you would be expecting the exam on Thursday, and it would not be a surprise. You can apply the same reasoning to rule out Monday-Wednesday, so you reach the conclusion that it would be impossible to have a surprise exam next week. Which makes it all the more surprising when your teacher gives the exam on Tuesday!

The surprise exam paradox demonstrates our pattern of reasoning when we try to anticipate a surprise. The puzzle of the paradox is due in part to the inevitable contradiction inherent in expecting a surprise. By definition, a surprise is unexpected. Attempts to anticipate surprising events therefore involve counterinductive reasoning. If I am expecting a surprise exam, each day without the exam establishes a regularity of observed instances, and I reason counterinductively that this regularity will not hold for unobserved instances. Though seemingly irrational, counterinduction is a rule of inference uniquely suited to anticipating an anomaly that you expect to occur in an otherwise uniform sequence.

My argument diminishes the force of Salmon’s counterinduction parody, whose success depends on the presupposition that counterinduction is not rational, therefore any argument
supporting it must be fallacious. Indeed, compared with our inductive practice, counterinduction has far less practical utility, and the rule circular justification for counterinduction strikes us as an obviously irrational argument. I nevertheless maintain that Salmon’s parody is not sufficiently strong to undermine rule circular justification, because the rule he chose is in fact useful in our practice of anticipating a surprise.

Yet the counterinductive reasoning employed in the surprise exam paradox must be misguided, because it supports the false conclusion that it would be impossible to have a surprise exam next week. As it happens, your teacher gives the exam on Tuesday, and you do feel surprised. I believe your surprise is partially due to the inevitable contradiction inherent in anticipating a surprise: when you expect a surprise then it is no longer surprising, so you stop expecting it, and then it surprises you. More central to the paradox is the misapplication of counterinduction to reason backwards in time. After you rule out Friday as a potential day for the exam, you subsequently eliminate the remaining days of the week, while keeping fixed your belief that the later days have been categorically ruled out. Potential exam days can only be definitively ruled out, if you actually go to school without having the exam. Each day this happens, you use counterinductive reasoning to rationally infer that the exam will happen on some future date. Yet when you are merely imagining the course of the week in your attempt to anticipate the exam, you cannot hold fixed your beliefs about the future chances of the exam happening later in the week, in order to form beliefs about the chance it might happen earlier in the week. The flaw in this reasoning pertains to the temporal asymmetry of chance, where past events can no longer be considered chancy.

The temporal asymmetry of chance poses a unique problem for counterinductive reasoning to anticipate a surprise, which does not arise in our standard inductive practice of predicting and explaining chance events. Counterinduction, as a rule of inference, also exhibits temporal asymmetry. A counterinductive inference depends on the established regularity of observed instances, from which you counterinductively infer that the regularity will not hold for unobserved instances. When anticipating a surprise, counterinductive reasoning can therefore only be applied to reason forward in time, according to the rule’s temporal asymmetry, as well as the temporal asymmetry of the chance that the surprise will occur.
In the previous section, I framed reasoning about objective chance within the broader context of our inductive practices. When reasoning about chance, our practical ability to discern probability relies on the frequencies we observe. Subjective credences follow from our patterns of observations, which in turn inform our credence that those patterns will continue. There is a lurking circularity here, which corresponds to the epistemic circularity characterizing our inductive practice. I will critique another philosopher’s attempt to dissolve this circularity in inductive justification. While I believe that the argument ultimately proves to be unsuccessful, the argument’s allusion to metacognition provides a deeper understanding of irrational feelings of surprise.

William P. Alston evaluates the potential for epistemic circularity to undermine justification of our basic belief-forming mechanisms. The most adequate concept of epistemic justification puts a reliability constraint on principles of justification. Reliability, in this context, refers to the tendency for a mechanism to produce true beliefs in the sorts of situations in which it normally functions. Cogent arguments that intend to justify claims of reliability about basic sources of belief tend to presuppose the reliability of the source in question. Epistemically circular arguments are therefore convincing only if one is already disposed to accept the reliability of a belief-forming mechanism. Alston argues that we have a misguided tendency to take the reliability of our basic sources of belief as self-evident, because of our strong inclination to accept such reliability without question. I will first critique an attempt to dissolve epistemic circularity in justifying our inductive practice, then I will apply the argument’s analysis of metacognition to better understand our feelings of surprise.

Alexander Jackson attempts to dissolve epistemic circularity in his argument justifying our inductive practice. Jackson reframes the problem of induction to focus on category-based induction, where the assumption of continued uniformity of some characteristic is localized to a specific context within space and time. Rather than centering the problem of induction around the presupposition of some general principle of the uniformity of nature, when inductively
reasoning from E to H (where E and H are defined according to the given specified context), we rely on the unjustified presupposition of ‘if E then H.’ Jackson argues that inductive inferences do not in fact rely on presupposition. Instead, when an inference stands with judging that E → H, the high degree to which the inference seems correct causes one to both judge that E H and reason from E to H. Jackson characterizes this as ‘cosupposition,’ where our belief in the proposition E → H is formed simultaneously with our belief that we are correct in inferring the next H. I understand Jackson’s argument as an attempt to collapse the epistemic timeline of inductive reasoning, where an inference and its corresponding inductive principle are cognitively formed in one fell swoop. Our inclination to infer H entails a corresponding acceptance of the proposition E → H.

I will present an example of category-based induction, defining E and H within a specific context. I leave my house on Monday morning at 10:00, and I see a black cat through my neighbor’s window. Leaving my house on Tuesday at 10:00, I see the cat again. I am fairly certain that it is the same black cat – its whiskers are distinctively scorched, just as I had observed the day before. I have the same experience on Wednesday and Thursday. There are various possible explanations for this pattern of observations: perhaps my neighbor feeds the cat every morning at 10:00, or perhaps my neighbor is a witch and put a hex on me. Having observed this explainable pattern of events, I make the inductive inference on Friday morning that I will see the cat when I leave at 10:00. I might even be a little surprised if I did not. Applying Jackson’s argument, my inference that the cat will be there on Friday is cosupposed with the principle, “if I leave my house at 10:00, I will see my neighbor’s black cat.”

Jackson attributes this process of cosupposition to a two-alternative forced choice, motivated by our metacognitive ‘feeling of rightness’ attached to our inferring H. A two-alternative forced choice presents two exhaustive, mutually exclusive alternatives, where logic demands acceptance of one proposition and rejection of the alternative. In this instance, the two alternatives are inferring H, versus refusing to judge that E → H. Jackson argues that the strong feeling of rightness attached to the representation of H does all the work in resolving this decision task, because there is no feeling of rightness attached to the judgment that E → H. Our belief that E → H is essentially along for the ride when we choose to infer H, because our
acceptance that $E \rightarrow H$ is demanded by our response to the two-alternative forced choice. Whatever makes reasoning from $E$ to $H$ rational, also makes it rational to judge that $E \rightarrow H$.

Jackson distinguishes this process of cosupposition from the unjustified presupposition of $E \rightarrow H$, because the formation of our belief that $E \rightarrow H$ is not epistemically prior to our inferring $H$. This notion of cosupposition introduces a distinctive kind of epistemic rationality constraint, where one must settle that $E \rightarrow H$ at the same epistemic point as one reasons from $E$ to $H$. Jackson argues that we question whether we have judged $E \rightarrow H$ in a rational way, because of the absence of an attached feeling of rightness. He advocates adding to our epistemic methods the knowledge by accompaniment involved in cosupposition, so that we can avoid the skeptical puzzle that arises from this blind spot in our psychology.

I see Jackson’s attempt to collapse the epistemic timeline of inductive reasoning as a promising approach, but the terms he outlines serve merely to disguise, rather than dissolve, the problem of circular reasoning involved in justifying our inductive practice. Jackson proposes this process of cosupposition in order to avoid the unjustified presupposition of inductive principles when forming inductive inferences. I however understand cosupposition as necessarily entailing circularity in reasoning, where Jackson’s denial of the epistemic priority of assuming $E \rightarrow H$ is incompatible with his argument for a feeling of rightness attached to our inferring $H$. In the example of seeing my neighbor’s cat, I had no feeling of rightness attached to the inference that I would see the cat on Monday morning. The inference I form on Friday followed from the pattern that I observed throughout the week. I formed the belief that, “if I leave my house at 10:00, I will see my neighbor’s black cat,” prior to making the inference that I would see the cat on Friday morning.

Metacognitive judgments induce feelings of rightness that guide us through philosophical reasoning and problem-solving. These judgments support foundationalism that shields us from skepticism, where a feeling of rightness can indicate the arrival at some basic or fundamental logical principle, beyond which we are incapable of reasoning. The feeling of rightness attached to our recognition of valid deductive arguments is necessary to avoid the Lewis Carroll problem. Our metacognitive judgments however, are not perfectly reliable, as evidenced by instances of cognitive or perceptual illusion. An example of the former would be our responses to the
cognitive reflection test, a task designed to measure a person’s tendency to override initially incorrect gut responses to simple problems, and engage in further reflection to find the correct answer. One question on the test is: *A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost?* A typical person’s initial gut response is that the ball must cost ten cents, but upon further reflection, she might come to realize that the ball actually costs five cents. The incorrect initial response is motivated by a sort of cognitive sleight of hand, similar to the visual sleight of hand involved in performing card tricks. A visual illusion relies on distraction, drawing our attention to features of the perceptual content that do not reliably guide us in forming justified beliefs. In the above case of cognitive sleight of hand, our distracted preoccupation with the difference in price between the bat and the ball prompts us to falsely attach a feeling of rightness to the belief that the bat must cost one dollar. The possibility of such cognitive or perceptual distraction is cause for us to doubt our feelings of rightness, which might attach to incorrect conclusions to problems that require further rational examination.

The feeling of rightness that Jackson argues attaches to our inferring H is a similar result of cognitive sleight of hand, because the feeling in fact depends on the epistemic priority of presupposing that $E \rightarrow H$. Past observations of correctly inferring H from E support the belief that $E \rightarrow H$, which attaches a feeling of rightness to inferring the next H. It would not be rational to allow an unjustified feeling of rightness to do the work in resolving a two-alternative forced choice in the way Jackson describes, because such feelings can attach to false conclusions. A feeling of rightness attached to an inductive inference demands investigation to support it, and further reflection would lead us to acknowledge the unjustified presupposition of granting epistemic priority to the inductive principle $E \rightarrow H$. Jackson’s attempt to avoid a circular justification of our inductive practice is therefore unsuccessful, because the presupposition of certain inductive principles motivates our feelings of rightness attached to corresponding inductive inferences.

I believe that Jackson nevertheless makes a valuable contribution to the project of justifying our basic belief-forming mechanisms, in his consideration of metacognitive judgments. There is potential for some cognitive distraction or sleight of hand to attach a feeling of rightness
to an unreliable inductive principle, just as irrational feelings of surprise push us toward unreliable alternative explanations about objective chance.

Metacognition refers to the capacity to monitor, and potentially control, one’s own cognitive states. Metacognitive judgments oversee object-level cognitive operations, where subjective beliefs and feelings guide our strategies for problem solving. These feelings are not perfectly reliable, as evidenced by the familiar metacognitive experience of the ‘tip-of-the-tongue state.’ This phenomenon of failing to retrieve a word from memory, combined with partial recall and the feeling that retrieval is imminent, indicates a discrepancy between subjective and objective knowing. Cognitive sleight of hand exploits this discrepancy, so that a subjective feeling of rightness attaches to objectively unjustified propositions.

Metacognitive monitoring is still generally accurate, posing an interesting puzzle of identifying the particular factors that might prompt us to form misguided metacognitive judgments. This problem is complicated by the phenomenal quality of our metacognitive feelings, which are associated with a sense of self-evidence. The cue-utilization view of metacognition postulates that metacognitive judgments are inferential in origin. On this view, the accuracy of a metacognitive judgment corresponds with the reliability of the cues and heuristics on which it is based. This theory distinguishes between experience-based and information-based judgments: the former being sheer subjective feelings and intuitions with perceptual-like quality, and the latter being reasoned cognitions grounded in a network of beliefs. Asher Koriat compares this distinction to that between reason and emotion, which can conflict and pull judgments and behavior in opposite directions. Information-based cues speak to the relationship between confidence and accuracy for metacognitive judgments, because a subject’s inaccurate preconceptions about a given cognitive ability would skew her corresponding metacognitive judgments.

Experience-based metacognitive judgments have the phenomenal quality of an immediate unexplained intuition, similar to that of perceptual experience. Cue-familiarity correlates with a feeling of knowing, and correspondingly, lack of familiarity can serve as a basis for determining that something is not known. In contrast with information-based judgments, which rely on an explicitly inferential process, experience-based judgments involve a two-step process. Some
heuristic triggers a sheer subjective experience, which can then serve as the basis for noetic judgments. While the potential inaccuracy of information-based metacognitive judgments corresponds with a subject’s inaccurate preconceptions, experience-based judgments can be led astray by a breakdown in this two-step process. Metacognitive experiences have implicit antecedents and explicit consequences, where the processes that take off from subjective experience generally do not have access to the cues that produced that experience.

Valerie A. Thompson explains metacognition in the context of a dual-processing theory. According to this interpretation, reasoning and decision making are accomplished by the joint action of two types of processes: System 1 and System 2. S1 processes are fast, automatic, and give rise to a highly contextualized representation of the given problem, whereas S2 processes are more deliberate and decontextualized. The heuristic responses of S1 are cued automatically by perceptual or cognitive input, so the origins of these impressions are not available to introspection. The nature of S1 processes to contextualize input may result in the omission of relevant information, which does not obviously appear salient in the given environment. Thompson focuses on questions surrounding the circumstances and extent to which S2 processes intervene to modify S1 responses.

Under Thompson’s interpretation, metacognitive processes provide an important link between heuristic and analytic processes. Metacognitive judgments are causally relevant to decisions regarding whether to stay with a current cognitive output or seek another. S2 operates on biased or incomplete representations generated by S1, only if there is compelling reason to scrutinize those representations. This illustrates the potential for breakdown in the two-step process described by Koriat, because S2 intervention is not reliably cued when modification is necessary. In cases of cognitive illusion, a heuristic response is generated with a strong intuition that the answer is correct, and the feeling of rightness reduces the probability of S2 intervention.

In contrast, when reasoning about objective chance, metacognitive monitoring is overly sensitive to irrational feelings of surprise. While a surprising chance outcome sometimes rationally motivates an alternative explanation about the chances, our metacognitive judgments produce a disproportionate number of false positives, where irrational feelings of surprise create the mere illusion of explanatory urgency. While we ought to second guess feelings of rightness
that attach too readily to false propositions, we should also second guess our feelings of surprise that push us toward alternative chance explanations. We cannot help but feel surprised at certain chance events, but we can gain practice resisting our instinctual tendency to let these feelings govern our explanatory practice.


