WHAT MAKES SOMETHING SURPRISING?

Dan Baras & Oded Na'aman

Abstract: Surprises are important in our everyday lives as well as in our scientific and philosophical theorizing—in psychology, information theory, cognitive-neuroscience, philosophy of science, and confirmation theory. Nevertheless, there is no satisfactory theory of what makes something surprising. It has long been acknowledged that not everything unexpected is surprising. The reader had no reason to expect that there will be exactly 190 words in this abstract and yet there is nothing surprising about this fact. We offer a novel theory that explains when and why an unexpected fact is surprising. We distinguish between descriptive and normative notions of what is surprising; clarify the sense in which surprising facts are unexpected; and, finally, develop and defend the significance account of surprise, according to which a fact is surprising to an agent if and to the extent that it is both unexpected and significant to the agent. Since a surprising fact can be significant to an agent in various ways—personal, moral, epistemic, and aesthetic—surprise is not merely or primarily epistemic. Fitting surprise reflects more than a person's view of what *is*; it reflects a person's view of what *is significant*.

It is sometimes taken for granted that surprise is called for by the unexpected—this is *the naive view of surprise*. However, counterexamples abound. Consider the exact distance between your

---

face and the computer screen (or the page) these words are written on. Consider the precise number of specks of dust on your desk. Look out the window and observe the exact shade of the sky or measure the velocity of the wind. It is safe to assume that whatever you find would not be expected. You might expect the distance between your face and the computer screen to be within a certain range, but you do not expect the exact distance; you might expect the sky to be some shade of blue, but you do not expect this precise shade; you might expect the wind to be strong, but you do not expect a specific wind velocity. However, it is also safe to assume that what you would find in each of these cases will be *unsurprising*. Since the unexpected and the surprising come apart, the naive view of surprise is false. What, then, calls for surprise?

Understanding surprise is important for various reasons. In our daily lives, surprises matter to us (Ely et al. 2015). Occasionally we enjoy surprises, look forward to them, and become indignant when they are spoiled; other times we dread them, are hurt by them, or try to avoid them as best we can. Moreover, surprises are like blushes in that they are self-revealing: they are involuntary reactions and yet they are cognitively informed by our view of the world around us and our attitudes toward it.

In addition, surprise plays an important role in many theoretical contexts.[1] Psychologists suggest that surprise is among a small set of basic human emotions present across cultures (Ekman & Cordaro 2011; Tomkins 1962). Information theorists evaluate new information by measuring how surprising it is (Shannon 1948; Baldi 2002). In cognitive science, a new paradigm is being explored and hotly debated, according to which minimizing surprise is the central task of human cognition (Clark 2013; 2018). In philosophy of science, it has been proposed that the distinction between experiments and simulations depends on potential

---

[1] For a recent survey of scientific research on surprise, see Munnich et al. (2019). We are not claiming that in all of these contexts the term "surprise" is being used in the same way. Clarifying what we mean by surprise and how it might be different from other uses will be the first task of this paper.

surprise (Morgan 2005; Currie 2018). In epistemology, some have proposed that we analyze beliefs in terms of potential surprise (Spohn 2012). Perhaps more than anywhere else, one finds theories of surprise in confirmation theory, because some believe that predicting surprises yields especially strong confirmation (Horwich 1982; White 2005). Finally, the notion of surprise is central to so-called 'fine-tuning arguments': some believe that the existence of a fine-tuned life bearing universe like ours is very surprising and as such it is a reason to believe in the existence of many other universes (Parfit 1998) or a god (White 2018).[2]

Despite the significance of the notion of surprise—in theory and practice alike—a satisfactory replacement for the naive view of surprise is yet to be found. It has long since been acknowledged that while surprising facts tend to be improbable, not everything improbable is surprising (Weaver 1948; Horwich 1982, 101). What makes a fact surprising, many theorists assume, is its improbability *plus* some additional feature. Thus, there are various attempts in the literature to identify the secret ingredient that makes the merely improbable surprising.

It has been suggested that surprises cast doubt on beliefs about the circumstances (Horwich 1982, 101); that surprises are not only token-improbable but kind-improbable as well (Schlesigner 1991, 102); and that surprises are contrary to expectations (Harker 2012). Common to these accounts is the view that surprise is determined merely by epistemic factors, e.g. probabilities, beliefs, and expectations.[3] We argue that these analyses are mistaken and previous discussions of surprise are fraught with confusion. Instead, we develop a novel theory of surprise. We argue that improbability is *not* a necessary condition for surprise and that fitting

---

[2] Within this literature, the notions of *surprising* and of *calling for explanation* are often treated as the same. We believe this is a mistake. A fact can be surprising without calling for explanation and can call for explanation without being surprising. For discussion, see Baras (forthcoming, sec. 3.1.1).

[3] New versions of epistemic accounts are still being explored. Recent proposals include Shogenji (forthcoming) and unpublished work by Roger White (presented at a workshop that we organized at the Hebrew University).

surprise is determined not only by epistemic factors. Put simply, we argue that surprise is fitting in response to what is *fittingly unexpected and significant*.[4]

In the first section, we clarify the question we aim to address. In the second, we consider in what sense must a fact be unexpected in order to be surprising and argue that to merit surprise it is necessary but not sufficient that a fact be fittingly unexpected. We move on, in the third section, to consider four distinct ways in which unexpectable facts can be surprising. In the fourth, we present our account—the significance account of what makes something surprising. In the fifth section, we explore some of its illuminating implications. In the sixth, we conclude.

I.   CLARIFYING THE QUESTION

Consider the following case:

> <u>Someone</u>**:** Someone won the lottery.

Is it surprising that someone won the lottery? The answer might seem obvious, but the question is multiply ambiguous. We therefore begin by clarifying the question.

'Surprise': Reaction vs. External Fact. The word 'surprise' may denote either a kind of affective reaction or that which merits this affective reaction; an occurrence *in the person* or a fact *in the world*, as it were. Thus, *my surprise* (in the first sense) is a reaction to *a surprise* (in the second sense).[5] To keep track of the distinction, we will use 'surprise' exclusively to denote the affective reaction and 'surprising' to describe that which merits surprise. So in asking whether it is *surprising* that someone won the lottery we are asking whether this event merits a *reaction of surprise*.

---

[4] Unfortunately, we lack here the space to argue individually against each of the previous theories of surprise. Instead, we develop our theory and point out our main objections to previous theories along the way.

[5] In noting this distinction, we follow Christopher Miller: "surprise denotes both an internal feeling and an external event" (Miller 2015, p. 5).

'Surprising': Descriptive vs. Normative. Saying that something is surprising is ambiguous between a descriptive and a normative claim: that it *causes* surprise or that it *justifies* it. We are concerned with what is surprising in the normative sense. We are not asking how people are likely to respond upon learning that someone won the lottery. Nor is our question about what normally induces surprise. In asking whether it is surprising that someone won the lottery we are asking about a certain kind of justification.

There are different kinds of justifications and here we focus on one specific kind, fittingness. When we believe a credible proposition, desire a desirable outcome, fear a fearsome bear, envy an enviable talent, or admire an admirable achievement, we have attitudes that are fitting to, or merited by, their objects. Having these attitudes in these cases might not be valuable, prudentially justified, or morally justified, and yet having these attitudes in these cases is fitting. Thus, in the current paper, we are concerned with fitting surprise. We offer a theory of what makes something a fitting object of surprise, or, in short, what makes something surprising.[6] This is what we mean by the normative sense of 'surprising'.

Conflating the normative and the descriptive senses of 'surprising' is a mistake, but it is a natural one to make.[7] While the two senses of 'surprising' are logically independent, they are

---

[6] For more on fittingness as a distinctive normative category see Howard (2018). For an influential paper on the fittingness of emotions as distinctive from moral and prudential justifications of emotions, see D'Arms and Jacobson (2000b).

[7] The two senses have often been conflated in research on surprise, both in and outside philosophy. Examples: Horwich (1982, 101) begins with the normative notion: 'we can be reasonably taken aback by…' but then, in the same paragraph, shifts to the descriptive notion: 'things...which don't surprise us' 'would surprise me.' Schlesinger (1991) and Harker (2012) repeatedly use descriptive terms, such as, 'will not surprise us' 'we are surprised by'. But given that the role they attribute to surprise is normative--i.e., determining epistemic support for competing theories--it is questionable whether the descriptive notion is really what they have in mind. Shogenji (forthcoming) says: 'it is not my goal to predict the degree of surprise... Surprise is a psychological phenomenon that resists simple generalization'. So what is his account an account of? He continues 'The analysis of surprise I propose aims to make paradigmatic cases of surprise understandable. It is informed by the way people commonly respond to paradigmatic cases of surprise.' This is still ambiguous between an explanation of what causes surprise in paradigmatic cases of surprise and an explanation of what makes surprise fitting in paradigmatic cases of fitting surprise.

In psychology there is a long line of research on factors that determine surprise. One way in which this research is conducted is by distributing questionnaires in which people are asked to rate how surprising some occurrence was or would be (Falk 1989; Maguire et al. 2011). As far as we can tell, these questionnaires do not clarify whether they ask participants how surprised they should be in the scenario, or how surprised they would be.

not unrelated. Intuitions and assumptions about what merits surprise are likely shaped, in part, by observations about what tends to elicit surprise. On the other hand, predictions of what might induce surprise are likely shaped, in part, by judgments about what merits surprise. Therefore, we should expect significant correlation between what we think merits surprise and what actually tends to induce surprise. In addition, once we understand what makes surprise fitting, we will have the tools to assess the fittingness of actual patterns of surprise, as uncovered by empirical research. Indeed, toward the end of the paper, we will use our normative account to shed light on some empirical findings about surprise.

Warranted Surprise vs. Fitting Surprise. Sometimes surprise merely *seems* fitting. Suppose you (think you) run into someone you know while trekking in a distant country. You have no reason to doubt your impression that this is indeed your acquaintance, so your surprise is warranted. But maybe, unbeknownst to you, your acquaintance has an indistinguishable doppelgänger. You think you ran into someone you know, but the person in front of you is a complete stranger. What you are surprised by does not obtain; you had sufficient but crucially misleading evidence to be surprised by it. Following D'Arms and Jacobson (2000a) and Scarantino and de Sousa (2021), we will say that in such cases your surprise is warranted but not fitting. Whether an attitude is warranted depends on whether the evidence justifies taking it to be fitting. Since norms of fit inform norms of warrant, we focus in this paper on norms of fitting surprise.

Surprise as an Emotion. We are interpreting the question 'is it surprising that someone won the lottery?' as a question about whether the fact that someone won the lottery merits a certain *emotion*. We note, however, that it is common to invoke surprise without implying an affective experience. Suppose that, while chatting next to the coffee machine, a colleague tells you she was surprised there was no traffic on the way to work today. You would probably not infer that your colleague experienced a certain emotion on the way to work, but merely that she expected

there to be traffic and discovered there was none. We often speak of surprise to denote our judgments and evaluations without implying anything about our actual affective reactions; we *borrow* the term to express our judgment and point of view. But suppose that later in the conversation your colleague tells you that her friends threw her a birthday party the other day: "Boy, was I surprised!" This time the report is clearly about her affective reaction, not only about her judgment of the fact in question. This is the *affective use* of 'surprise', on which we focus in this paper.

Partial motivation for this focus comes from the fact that often the non-affective use of 'surprise' is parasitic on the affective notion of surprise. When your colleague says she was surprised that there was no traffic on the way to work, she might be implying that it would have been fitting for her to experience surprise even if in fact she did not. We find the same phenomenon elsewhere: 'I'm happy to announce...', 'we regret to inform you…', 'I hope you understand...'—we often express our evaluations of various facts by invoking the emotions they merit without implying that these emotions actually obtain.

We are not alone in treating surprise as an emotion (Mellers et al. 2013; Munnich et al. 2019). Surprise seems to have features that are generally held to be characteristic of emotions. Surprise has a distinctive phenomenology—it *feels* a certain way; it is an intentional attitude, which is to say that it is *about* something or it is *directed* toward it;[8] and it can be fitting or correct with regard to what it is about. These are typical characteristics of emotions (Deonna & Teroni 2012; Scarantino & de Sousa 2021). Moreover, insofar as there are 'basic emotions'—emotions that are innate and common across cultures and societies—surprise is normally considered to be one of them (Ekman & Cordaro 2011; Tomkins 1962).

---

[8] For a discussion of the intentionality of emotions, see Goldie (2002, 16–28).

Although we treat surprise as an emotion, we also believe surprise has distinctive features among the emotions. One such feature is that emotions typically have a certain valence. Sadness, fear and disgust have a negative valence, while joy, love and amusement have a positive valence. Surprise, however, can have either positive or negative valence. One can be positively surprised by a surprise party as well as negatively surprised by a surprise attack. An alternative proposal is that surprise is neutral but is often tied to other, valenced emotions.[9] According to this view, for instance, when a person is 'positively surprised' by winning the lottery, she is *neutrally* surprised by something that she is very favorable toward. Either interpretation of the phenomenon would suit our purposes here. The emotion of surprise itself can either be positive or negative, or it is neutral but often coupled with either positive or negative attitudes.

Surprise is unique among the emotions in another way. Anger, disgust, or fear, for example, may fittingly persist for some time because the conditions that merit their occurrence can persist.[10] By contrast, the very occurrence of surprise implies that the conditions that merited it have changed. Once a person is fittingly surprised by a fact, she comes to believe it obtains and cannot be fittingly surprised by this fact *again*, even if the initial surprise fittingly lingers for a while.[11]

---

[9] We thank Yair Levy for the suggestion.

[10] Though, occasionally, the past existence of these emotions can undermine the fittingness of their future existence, see Na'aman (2019).

[11] Following Meredith Osmond, Christopher Miller maintains that this feature of surprise is expressed in language:

> We say that we are surprised *by* rather than *with* something … the distinction is based on the duration of response: one feels angry *with* someone because the emotion can be sustained, and its embodied cause continues to exist. But one is surprised only by the sudden or unexpected, after which the emotion vanishes or turns into something else" (Miller 2015, 5).

II.   EXPECTATIONS

We can now make the question about <u>Someone</u> more precise: does the fact that someone won the lottery make fitting an emotional reaction of surprise? The answer is, clearly, 'no'. Assuming, of course, that the lottery in question is one where someone is guaranteed to win. Under this assumption, it is not surprising that someone won the lottery because it is to be expected. The example suggests that, in general, surprise is fitting only in response to what is unexpected. But what does "unexpected" mean in this context?

Fitting vs. Actual Expectations. As before, it is important to distinguish between descriptive and normative senses of the notion we are considering—that is, between what is *actually* unexpected and what is *fittingly* unexpected. We suggest that expecting X is fitting just in case X is expectable or predictable. Furthermore, and perhaps less obviously, we assume that a lack of expectation is a state that can be unfitting. In support of this assumption, we note that a lack of expectation often strikes us as a failure. Driving through a residential neighborhood, it is unfitting not to expect pedestrians; when dark clouds cover the sky and thunders are heard, it is unfitting not to expect rain. Generally, when that which one does not expect is expectable, lack of expectation is unfitting ("how could you not see it coming?!").[12]

---

[12] Although there is little writing on the subject, it is occasionally claimed that a lack of an attitude cannot be fitting. A possible line of reasoning behind this thought was suggested to us by an anonymous reviewer: A fitting attitude is fitting to the object it is about, but a lack of an attitude is not about anything so it cannot be fitting to anything. If this reasoning is correct, then it seems like we are making a mistake in claiming that for a fact to be surprising it must be fittingly unexpected. The first thing to note about this is that, as we discuss shortly, often we do not expect p by expecting not-p, at least implicitly. In such cases the problem doesn't arise because expecting not-p is an attitude and as such can be fitting. What about cases where neither p nor its negation is expected? One simple way to get around the problem is to posit that in such cases, when we say that it would be fitting not to expect p, all we mean is that expecting p would be unfitting. There are other theoretical possibilities. For instance, some might argue that a lack of attitude can be fit-evaluable after all or that a lack of expectation is an attitude in the same way that suspension of belief is, at least according to recent authors (Friedman 2013). We take this to be an interesting question and fertile ground for future research. (The authors found themselves disagreeing about various theoretical issues here). At any rate, we are more committed to our judgments about the examples, showing that it is unfitting to be surprised by clearly expectable facts, than we are to a particular view about whether and how a lack of attitude can be fitting.

Descriptive and normative unexpectedness can come apart. Suppose it is your birthday and you're walking back home from work. Outside your apartment building, you notice your friends' bicycles parked side by side along the wall. Let us suppose that on the basis of this fact it is fitting for you to infer and unfitting not to infer that your friends are in your apartment, waiting for you. That is, you should expect a surprise party.[13] Suppose however that you do not draw this inference and therefore do not expect a surprise party. You would likely then be surprised by seeing all your friends as you walk through the door. But would your surprise be fitting in this case? That is, can surprise be fitting in response to something expectable?

The opposite can also happen. Suppose you have absolutely no reason to suspect that a surprise party awaits you. However, as you walk towards your house, you imagine a surprise party and wishfully form the expectation that your friends are waiting behind your apartment door with balloons and a big cake. It turns out that you're right: they *are* waiting for you. Because you formed the expectation, you are not surprised to see them in your apartment. Is your lack of surprise fitting? Or would it be fitting to be surprised despite the fact that you actually expected the party because the party was unexpectable? Obviously, what determines whether one is *actually* surprised is typically one's actual expectations. But what determines whether one is *fittingly* surprised: actual expectations or fitting expectations, or maybe both?

Three possible views suggest themselves. (1) On the first view, a fact X is surprising to A only if A did not *actually* expect it, but it is not necessary that this lack of expectation is fitting; (2) on the second view, a fact X is surprising to A only if it *would have been fitting* for A not to expect it, but it is not necessary that A actually doesn't expect X; (3) on the third, a fact X is surprising to A only if A *fittingly* did not *actually* expect it. All three views allow that a person's

---

[13] Henceforth, when expectation is fitting we will sometimes say that one should expect.

expectations can be unfitting, what the views disagree on is whether the flaw in the agent's expectations implicates the fittingness of the agent's surprise.

We maintain that the third view is the correct one. View (1) is implausible because in cases in which a fact is entirely expectable, surprise seems unfitting even if the agent did not expect the fact. For instance, if you told us you are coming to a party, and yet we are surprised to see you there, our surprise seems unfitting because we should have expected you to be there (i.e., it was unfitting for us not to expect to see you there). The fact that we didn't expect you to come might explain our surprise but it wouldn't make it fitting. View (2) is implausible because surprise does not seem fitting in cases in which an agent unfittingly expects a fact. Suppose Guy wakes up this morning with an inexplicable hunch that he will run into his childhood friend, Abigail, whom he hasn't seen in a long time. He believes his hunch is an omen and forms an expectation that he will run into Abigail. Guy has no *good reason* to expect this, but he does. And, indeed, he runs into Abigail on the subway later in the day. We suggest that given that Guy had no reason to expect to run into Abigail, surprise would have been fitting if he didn't expect to run into her. But, we also suggest that, given that he *did* expect (though without good reason and thus unfittingly) to run into her, it would *not* be fitting for him to react with surprise. Such examples suggest that actual expectations matter for surprisingness. View (3) captures the way in which actual expectations and fitting expectations combine to determine whether a fact is surprising. On our account, whether a fact is surprising to an agent depends on whether it is both unexpectable and unexpected; or, in other words, whether it is *fittingly and actually unexpected* by the agent.[14]

---

[14]Although we maintain that a surprising fact must be both unexpectable and unexpected, we often mention only one of these conditions for purposes of simplicity. However, by using the one we don't mean to exclude the other.

Binary vs. Graded Notions. As with many other attitudes, we have two ways of talking about both fitting surprise and the corresponding property of being surprising: we can treat them as binary or as graded. We often say that such and such is surprising, or unsurprising, while other times we appeal to grades of surprisingness, saying that something is very surprising, or a bit surprising, or more surprising than something else. The same is true of the way we speak and think of expectations. In this respect, the two notions, surprise and expectation, map nicely on to each other. Just as a fact must be unexpectable to be surprising, the more unexpectable it is the more surprising it is.

This raises the question, what precisely is the relationship between the binary and the graded notions of 'surprising'? For example, is something surprising *simpliciter* iff it is surprising to a degree above a certain threshold? The same proposal might be made with regard to what is unexpectable: something is unexpectable *simpliciter* iff it is expectable to a degree below a certain threshold. These proposals follow a natural way of analyzing belief in terms of credence by employing a so-called "threshold account". But we do not think a simple threshold account holds promise for the relation between belief and credence and for similar reasons we do not think it would prove successful as an analysis of surprise and expectation.[15] We leave the question open, but note that a plausible view of the relation between credence and belief is likely to shed light on the relation between the graded and binary notions of 'surprising'.

Expectations vs. Probabilities. A number of philosophers have argued that low prior probability is a necessary condition for surprise. For example, Paul Horwich says that "the improbability of an event is not sufficient – but it does seem necessary" (1982, 101); Neil Manson and Michael Thrush say that "[i]mprobability…seems to be … a necessary condition

---

[15] The threshold account of the relation between belief and credence is often called the 'Lockean thesis'. For a survey of arguments against it, see Jackson (2020, sec. 3.2).

for an event's being surprising" (2003, 80); and David Harker says that "surprises are epistemically improbable" (2012, 253).

These authors may be right about the relation between low prior probability and surprising *events*, but events are not the only things that can be surprising. For instance, mathematical results can be genuinely surprising, despite having probability 1 according to standard probability theory. Googling "surprising mathematical facts" yields lists of such results. For example, the following fact:

$$1/89=0.0\underline{1}+0.00\underline{1}+0.000\underline{2}+0.0000\underline{3}+0.000000\underline{5}+0.0000000\underline{8}+\ldots$$

How should the three dots be filled in? Notice that 1,1,2,3,5,8 are the first numbers in the Fibonacci Series. If you keep adding a zero and the next number in the Fibonacci Series until infinity, it all adds up to exactly $1/89$.[16] An account of fitting surprise should explain how mathematical results can be surprising.

Thinking in terms of fitting expectations seems quite natural here. Given that we are not logically omniscient, it is fitting not to expect any of the surprising mathematical results prior to learning about them. In addition, mathematical training and comprehension can affect our fitting expectations. Thus, on our account, surprisingness is determined by fitting expectations, not by probabilities.

Now, some people react to this argument by claiming that there is an epistemic notion of probability that means the same as justified expectation. To them we say: if this is what you mean by probability, then our differences are merely verbal, and we agree that "low probability" as you interpret the term is necessary for surprise. But since many people think of

---

[16] A rich repository of examples of surprising mathematical results is Marc Lange's work on explanation in mathematics (Lange 2016, Chapters 7–8). Lange's concern is with explanation, not with surprise, but his examples are mostly of mathematical results that are, among other things, surprising.

probability, even epistemic probability, as something that must conform to Kolmogorov's (1950) three axioms of probability, and all standard interpretations of those axioms imply that necessary facts have high probability, we prefer using a different term: "expectation."

Occurrent vs. Dispositional Expectations. Outside our office is a botanical garden. We walk through it every day. Along the way there is an Arbutus tree. Since we pass by it every day, as we get closer to its location, we expect to see it there. This expectation might be *occurrent*, meaning that it actually occurs to us as we walk by. However, sometimes, when walking to our office, our mind drifts in thought and instead of thinking of the upcoming Arbutus tree we think about a philosophical problem. We do not give any thought to the fact that we are approaching the Arbutus tree. Do we, in those moments, expect the tree to be there? There is a sense in which we do: we have a dispositional expectation. If you interrupted our stream of thought and asked: "do you expect an Arbutus tree to appear around the corner?" we would say "yes!" without hesitation. Although the thought had not occurred to us at the moment, it had occurred in the past and is easily retrieved from memory.

Sometimes expectations can be even less explicit. Suppose one day we walk around the garden and come across a pink elephant standing right in the middle of our path. We should be extremely surprised despite the fact that, prior to writing this passage, the idea of a pink elephant standing there never occurred to us, nor did we have reason to consider it. Fitting expectations that determine fitting surprise need not be explicit or occurrent; they can be dispositional or occur at a subconscious level.[17]

Not Expect vs. Expect Not. When we say that a surprising fact must be fittingly unexpected should we mean that one must fittingly not expect it or that one must fittingly expect it not to

---

[17] Is an expectation a belief? We are inclined to think of belief as something more explicit than an expectation. However, this is not common ground and, so long as we have a general grip on what an expectation is, we need not take a stand on this issue for the purposes of this paper.

be? The two meanings come apart if there are cases in which one has no expectation one way or another. Such cases would illustrate that expecting not-p is one way of not expecting p, but not the only way.

Consider: Do we expect to find a pink elephant in the botanical garden? If the question is only about explicit expectations, then the answer is that until this very moment we had no expectation one way or another. The thought of a pink elephant never crossed our minds. But, as we already argued, fitting explicit expectations are not the only kind of expectations that determine surprise. Once we include our fitting implicit or dispositional expectations, the case should be described as one in which we expect that no pink elephant will be in the garden.

A different kind of case is when one has roughly equal reason to expect p and to expect not-p. In such a case, one should not have any expectation, explicit or implicit, one way or another. Given such a situation, can finding out that p (or that not-p) be surprising? We think either result can be somewhat surprising. Imagine you bought half the tickets of a high-stakes lottery. Would it be fitting to be surprised that you won? Obviously, you should not be very surprised because there was a 50-50 chance. But perhaps some surprise is fitting: you won a great deal of money which you could have lost just as easily. This case suggests that not expecting p need not imply expecting not p.[18] So, the necessary condition for fitting-surprise is that you fittingly do not expect the fact, not that you fittingly expect its negation.

To sum up our lessons from Someone:

---

[18] There might be a third kind of case in which one has no expectation whatsoever, not even implicit, regarding p. In this third case one lacks expectation about p not because p and not-p are equally expectable, but because people don't have expectations about every possible fact or proposition and this person happens to lack an expectation about p. See footnote 12, where we discuss possible interpretations of the idea that a lack of attitude can be fitting.

UNEXPECTED: An agent's surprise is fitting only in response to a fact the agent fittingly does not expect.

UNEXPECTED strikes us as a necessary condition for fitting surprise, but its satisfaction is not sufficient for fitting surprise. Consider another case:

Robinsons: John and Lisa Robinson from Tennessee won $533 million in the lottery.[19]

Robinsons satisfies UNEXPECTED. It was expectable that someone would win the lottery but it was certainly unexpectable that the Robinsons would win the lottery. The odds were 292 million to one. And yet, we submit, Robinsons is not surprising. It was expectable that some particular person or persons would win; why should it matter whether the winners are the Robinsons, the Jacobsons, or the Simpsons? The fact that this particular couple won the lottery is as surprising as the fact that there are 10,581 leaves on the tree. Both facts are fittingly unexpected but neither is surprising. So UNEXPECTED is a necessary but not sufficient condition for fitting surprise.

It might seem possible to object to our last claim and insist that the Robinsons' win is surprising. The objector might appeal to a view that we noted above (section I), namely, that surprise is neutral: surprise has no valence in itself but is occasionally and not necessarily tied to other, valenced emotions, such as joy or sadness. It is true, the objector might say, that Robinsons does not merit joy or gladness, but it nevertheless merits surprise, given that the Robinsons' win was highly unexpectable. Once we distinguish fitting surprise from fitting joy, the objector claims, we can see that the Robinsons' win merits surprise even if (and for whom) it does not merit joy.

---

[19] Associated Press, "Lottery officials verify Tennessee couple wins third of $1.6-billion Powerball jackpot", *LA Times*, January 15, 2016. From here onward, we move to imaginary variants of this example.

We disagree: Robinsons does not merit surprise quite independently of whether it merits other reactions. Note, first, that even if surprise is taken to be neutral, it is not to be confused with a non-affective judgment that a fact was unexpectable; surprise is an affective, emotional reaction. Second, the idea that surprise is neutral implies that we can, in principle, feel surprised by some things without feeling positively or negatively about them. So a great but neutral surprise is intensely felt. However, we find it implausible that the fact there are 10,581 leaves on the tree merits an intense emotion of surprise and, similarly, we find implausible that the Robinsons' highly unexpectable win merits an intense emotion of surprise, whether valenced or not. Next, we consider what could make the Robinsons' big win surprising after all.

III.   FOUR KINDS OF SURPRISES

To understand what makes some unexpected facts surprising, we consider in this section four examples of facts that, unlike Someone and Robinsons, would be surprising. These four examples represent four kinds of surprises. They will lead the way to the unified account that follows in the next section.

Consider first, Friends:

> Friends: John and Lisa Robinson from Tennessee won $533 million in the lottery and we are close friends of the couple, or family members, or we *are* the Robinsons.

The Robinsons' win is not surprising to complete strangers, but in Friends we imagine that the Robinsons are not just a random couple, but our friends, or family members, or that we are the Robinsons. In these cases, the fact that they won the big prize seems very surprising to *us*. It would be odd, to say the least, if friends and family of the Robinsons remained indifferent to the news of their big win; the Robinsons' win merits their surprise. However, insofar as the Robinsons are complete strangers to you, then *for you* their win is unsurprising. What might

explain the change from <u>Robinsons</u> to <u>Friends</u>? Why is the fact that the Robinsons won surprising to them, their friends, and their family, but not for complete strangers?

One might first be tempted to answer that the probability that our friends or family will win the big prize is much smaller than the probability that some random couple we don't know will win. This is true. But the probability that a specific random couple we don't know—say, the Jacobsons—will win is also much smaller than the probability that some random couple we don't know will win. We are considering the contrast between the unexpectable and unsurprising win of a specific couple we don't know and the similarly unexpectable yet surprising win of a specific couple who are our friends or family.

We propose, instead, that what makes the unexpectable fact in <u>Friends</u> surprising is that it has a *significant personal impact*, that is, a significant impact on something they value and care about. In <u>Robinsons</u>, the Robinsons' win does not matter one way or another to the people and things we value, but in <u>Friends</u> it does. Indeed, this explanation fits a wide range of examples. It is fitting for you to be surprised by a surprise party for you, but unfitting to be surprised by a stranger's surprise party. You might be fittingly and positively surprised if you are unexpectedly offered an amazing job, and negatively surprised if you are unexpectedly fired from a job you love. But it would not be fitting to be surprised if you discovered that the total number of leaves in your campus is 641,959,345, as unexpectable as this might be.

Personal impact is, as the phrase suggests, highly agent-relative. In contrast, some facts can have an impact that should be important to all people. Consider another variant of the case:

> **Vaccine**: John and Lisa Robinson from Tennessee won $533 million in the lottery. Before buying their ticket, the Robinsons vowed that if they win they will contribute the prize money to the search for a vaccine for COVID-19. Thanks to their contribution, a vaccine has been found.

In this case, the fact the Robinsons won has a significant impact on people and things you care about. After all, the global pandemic has touched most of our lives and a successful vaccine is the key to ending it. However, we submit that even without having a personal impact on the agent whose surprise is under consideration, the fact that the Robinsons won is surprising in Vaccine. We propose that this is because of the great *moral impact* of their win. Similarly, if global emissions of greenhouse gasses drop to net zero by 2030, or if global hunger suddenly triples, everyone's surprise would be fitting. The former fact would merit positive surprise and the latter negative surprise. Moral surprise, or moral shock, can be fitting when one comes across grave wrongdoing or an unpredictable situation in which one morally ought to help a person in need, as well as when unexpectable moral blessings occur, as in Vaccine. A distinctive feature of moral surprisingness is that, unlike personal surprisingness, it tends to be agent-neutral.

The ideas of personal impact and moral significance form only a partial explanation of what makes facts surprising. These explanations do not apply to all cases of fitting surprise. Consider yet another variant of our lottery example:

> **Nostradamus**: John and Lisa Robinson from Tennessee won $533 million in the lottery. Hundreds of years earlier, Nostradamus predicted that a couple who go by the same name will win this precise sum of money.

In this example, we stipulate that nothing we value will change significantly as a result of the Robinsons' win. The people we love and the activities we cherish would not be significantly affected by the scenario in Nostradamus. We also believe that Nostradamus doesn't present us with a morally significant fact; unlike Vaccine, we are not surprised here by something of moral importance. And yet, if Nostradamus predicted the Robinsons' win their win would be very surprising. Why?

Intuitively, the fact of the Robinsons' win in <u>Nostradamus</u> is surprising because it runs contrary to some deeply held beliefs. More precisely, it is a reason to revise some such beliefs, such as the belief that humans do not have supernatural powers and are not capable of making such predictions. We therefore suggest that the Robinsons' win in <u>Nostradamus</u> is surprising because of its *significant epistemic impact*. That is, the Robinsons' win is surprising in this case because it constitutes a reason to make significant revisions in our epistemic attitudes. This is epistemic surprisingness.

Epistemic surprisingness explains a wide range of examples. A coin landing many times HTHTHTHT... consistently, is surprising in part because it suggests that, contrary to initial appearances, the coin tosses are controlled by some invisible mechanism.[20] The history of science is stocked with examples of epistemic surprises, whereby an unexpectable observation surprises a scientist because it challenges previously held theories and prompts a new theory that changes what we believe about the world.

Our final example before we offer a unified account of fitting surprise, is the following case.

> **5-3-3**: John and Lisa Robinson from Tennessee won $533 million in the lottery. It was their 53rd anniversary, the 533rd ticket they ever bought, filled out at 5:33PM.

Even if we stipulate that the conjunction of facts in <u>5-3-3</u> has no significant impact on anything we value, no moral significance, and no significant epistemic impact, it still seems to be surprising. Examples of this sort are easy to contrive. Suppose, for instance, that a coin you know to be fair lands heads time after time after time. Or suppose you look up into the sky and see a cloud shaped like a dolphin. In these pure coincidence cases, there is neither significant impact on valuing nor significant epistemic impact because the fact in question is isolated from

---

[20] In fact, we think that surprisingness in this example can be explained in another way. The result of the tosses would be surprising even in a context in which the tosses are known to be fair. We will get to this shortly.

anything one values and doesn't matter one way or another to what one believes (that is, over and above its relevance to the prior lack of expectation of the fact itself.)

We suggest that what makes 5-3-3 surprising is its *substantive aesthetic quality*, and that the same is true of analogous cases. The fact the number '533' repeats in different variations in the details of the Robinsons' win is aesthetically striking. We notice this alignment the same way we notice a striking symmetry, a steady growth or decline, or a group of items that together make a series. Whether or not these facts have any further significance, their aesthetic quality makes them jump out at us, so to speak, and surprise us when unexpected. By contrast, there is nothing aesthetically noteworthy about a messy sequence of heads and tails or a randomly shaped cloud.

Our explanation of aesthetic surprises makes very minimal assumptions about aesthetic properties. We assume that facts can have aesthetic properties. We further assume that these properties imply that some facts are aesthetically boring and others interesting. Or, to put it differently, some facts are aesthetically neutral and others have a substantive aesthetic quality, positive or negative.

Friends, Vaccine, Nostradamus, and 5-3-3 suggest four possible explanations for fitting surprise. An agent's surprise is fitting in response to a fact the agent fittingly does not expect if it either (a) significantly impacts something the agent values, or (b) is morally significant, or (c) has a significant epistemic impact on the agent, or (d) has a substantive aesthetic quality. In the next section we turn to the question of whether there is a more general account that unifies these four explanations.

IV. THE SIGNIFICANCE ACCOUNT

What do impact on valuing, moral significance, epistemic impact, and substantive aesthetic quality have in common? We suggest that these are all ways in which a fact can be *significant* to an agent. We therefore propose that an agent's surprise is fitting in response to a fact that the agent fittingly does not expect and that is significant to the agent.[21]

Personal, moral, epistemic, and aesthetic significance are notions we can recognize quite apart from our judgments about surprisingness. To see this, imagine scenarios in which you encounter a fact that you *fittingly expect*. In such scenarios, surprise would not be fitting. And yet, it is still generally easy to judge whether the fact in question has significance or not. These judgments tend to fit nicely with judgments about whether the same fact would be surprising were it not expected. Suppose, for example, that you realize that your friends are organizing a surprise party for you. When you walk in the door and your friends yell "surprise!" you're not surprised because you expect the party. Nonetheless, the fact your friends threw a party for you is significant to you.

Significance is Graded. As we noted in our discussion of expectations, fitting surprise is graded. We saw that one factor that influences the grade of surprise is the grade of expectability. Winning a thousand dollars in a 100-ticket lottery is less surprising than winning the same sum in a 1,000,000 ticket lottery because the latter is less expectable. If the significance account of surprise is correct, then the grade of significance will also influence the grade of fitting surprise. This result seems to align with considered judgments about cases. Compare winning 1$ in a

---

[21] Interestingly, there is some empirical research that suggests that people tend to be more surprised by outcomes that they consider more significant (Gendolla 1997). For example, participants were told a story about a student failing an exam and were asked whether that student was surprised. Some participants were told that the exam was important to the student, others were told that it was unimportant. The former group expected the student to be significantly more surprised than the latter. We don't consider this as confirmation for our theory, since, as explained above, such experiments examine descriptive aspects of surprise whereas our thesis is about fitting surprise. Rather, it shows that there is some correlation between when we think surprise is fitting and when we expect people to be surprised.

lottery with winning 1,000,000$ in a lottery. Even if the odds of winning in each of the two lotteries are the same, you would probably be much more surprised in the latter case than in the former, and fittingly so.[22]

We therefore arrive at the following proposal:

> THE SIGNIFICANCE ACCOUNT: An agent's surprise by a fact is fitting if and to the extent that the fact is fittingly unexpected by that agent and significant to the agent.[23]

Descriptive vs. Normative Significance. To say that some fact (or object or event) is *significant to* an agent, is, once again, ambiguous between a descriptive and a normative reading. According to the descriptive reading, to say that a fact is significant to an agent is to refer to a certain *attitude of the agent*, an attitude that attributes significance to the fact in question. The attitude might attribute significance *de jure* by having as its content *that* the fact is significant; or the attitude might attribute significance *de facto* by treating the fact *as* significant (we will return to this distinction shortly). Alternatively, on the normative reading, to say that a fact is significant to an agent is to refer to *a normative property of the fact*, not to attitudes of the agent.

The two senses of significance can come apart. The existence of racial discrimination in our neighborhood is (normatively) significant to everybody in the neighborhood, even if it is (descriptively) significant only to some. However, sometimes normative significance partly depends on an agent's attribution of significance. For example, the significance of

---

[22] An anonymous referee suggests that the difference between the cases might only be in the grade of one's happiness or excitement, not in the grade of one's surprise. This proposal mirrors the proposal we discuss in Section I, according to which surprise might lack valence and therefore be distinguished from other, valenced emotions. If two facts are equally unexpectable, the thought goes, they are equally surprising, regardless of whether one merits great joy and the other little to none. Our response here is the same as before: even if surprise is not valenced, insofar as it is a felt emotion it seems unfitting that it be felt to the same extent when one wins 1$ and when one wins 1,000,000$.

[23] More precisely, though less elegantly, what we mean is to specify a necessary, sufficient and explanatory condition: An agent's surprise by a fact is fitting *if* and *only if* and *to the extent that* and *because* the fact is fittingly unexpected by that agent and significant to the agent.

birdwatching for Jonathan Franzen plausibly depends not only on the objective value of the activity but also on Franzen's significance-attributing attitudes toward birdwatching.

Which kind of significance determines fitting surprise, normative or descriptive? Three possible views suggest themselves, and, not-coincidentally, they are similar in structure to the three possible views we considered in our discussion of what is unexpected (section III). (1) On the first view, a fact X is surprising to A only if A attributes significance to X (X is descriptively significant to A), but it is not necessary that X is normatively significant to A; (2) on the second view, a fact X is surprising to A only if X is normatively significant to A, but it is not necessary that A attributes significance to X ; (3) on the third, a fact X is surprising to A only if A both attributes significance to X and X is normatively significant to X (X is both descriptively and normatively significant to A). As with *unexpectedness*, here, too, we opt for (3). To explain why, we need to say more about descriptive and normative significance.

First, to clarify the notion of significance-attributing attitudes, we draw on an influential paper by Samuel Scheffler. Scheffler distinguishes between recognizing something as *valuable*, or believing it has value, and *valuing* it. A person can recognize a great many things as being of value and yet actually value only a small subset of them (Scheffler 2010, 21). For example, Scheffler admits he recognizes the value of birdwatching and the study of Bulgarian history but does not value these activities. According to Scheffler, valuing X involves being susceptible to experiencing a range of context-dependent emotions with regard to the object, a disposition to experience these emotions as merited, and a disposition to treat certain considerations related to the object as reasons for action in certain contexts (Scheffler 2010, 29).

Without endorsing Scheffler's account wholesale, we wish to take from it the idea that valuing X involves a certain emotional and practical vulnerability to X and a commitment to seeing this vulnerability as fitting to its object. We take this to be a characterization of what it means

(in our present context) for an agent to attribute significance to X. Thus, if we spend a great deal of our time counting blades of grass and fussing over getting the number right, we would seem to be attributing significance to this activity. In this case you would be correct to assume that we are committed to the judgment that counting blades of grass is worthy of our emotional and practical engagement with it, and you may reasonably question our judgment.[24] We propose that X is descriptively significant to an agent when the agent values X.

Now let's clarify what is the normative property of being significant to an agent. Some things are normatively significant or important to you partly because you treat them as such and others are normatively significant or important to you regardless of whether you treat them as such (Frankfurt 1982). If Russian formalism and your friend Fred are significant to you this is partly because you care about them, but keeping a reasonably healthy diet or having a moral compass are important for you independently of whether you care about being healthy and avoiding wrongdoing. That said, even when caring about something is part of what makes something significant to you, what you care about must be worthy of the significance you attribute to it.

Our view is that what merits surprise is a fittingly unexpected fact that significantly bears on something the agent *fittingly values*. So what is required is both that the agent actually attributes significance of some kind to the fact *and* that this attribution is fitting. Consider, first, unmerited valuing. If an unexpectable thunderstorm prevents us from going outside to count blades of grass and this activity is one we value, then it significantly impacts something we value—the activity of counting blades of grass. We might be surprised, and our surprise might cohere with our judgment that counting blades of grass is significant—i.e., worth valuing—but our surprise would not be fitting because counting blades of grass is *not* worth valuing.

---

[24] While Scheffler argues that *valuing X* involves *believing X valuable*, it is possible to hold that one can value X without having a belief about X's value. We don't need to take a stand on the matter here. This is why we distinguish above between *de jure* and *de facto* attributions of significance to X. The former involves a belief in the value of X while the latter does not.

Similarly, if we are surprised to find that there are 10,581 leaves on the tree because this number of leaves, we believe, indicates that we have a bright future ahead of us, then our surprise is not fitting. Our surprise is not fitting, not because we should not value our future, but because the number of leaves on the tree does not in fact impact our future or indicate anything about it.

Second, consider optional valuing. A fact is not surprising if it impacts something we *may* fittingly value but we do not in fact value it. The unexpectable weather may be perfect for birdwatching, which we may fittingly value, but since we do not in fact value it the fact that the weather is perfect for birdwatching is not surprising to us. Third, we submit that a fact is not surprising if it impacts something one *ought to value but doesn't*. Even if a person ought to care about her brother, if she doesn't care about him then the fact he won the lottery or had a car accident, though unexpectable, would not be surprising to her. Not only would she not be surprised as a matter of fact, her surprise would not be fitting given that she doesn't attribute to her brother the significance she ought to attribute to him. To be sure, she would be criticizable for her lack of concern for her brother, but given her lack of concern her surprise would not be fitting.

The account of significance as *fitting attribution of significance* applies not only to personal and moral significance, it applies to epistemic and aesthetic significance as well. Not every epistemic change is significant to an agent. Learning that some particular ticket is the winning ticket is not in itself epistemically significant, though it involves an epistemic change. By contrast, learning that Nostradamus predicted correctly who will win the lottery is significant because it is a reason to change our beliefs about laws of nature and human abilities. But what is it about the latter epistemic impact that makes it significant?

We propose that the epistemic significance of an unexpectable fact be understood as its epistemic relevance to something the agent fittingly values. Learning new facts about the forefathers of the Robinsons would not be epistemically significant to strangers, but (arguably) it would be epistemically significant to the Robinsons and their relatives. Certain mathematical discoveries might be epistemically significant only to those who value math. Some facts might have an epistemic impact that is relevant to something everyone fittingly values, or they might be epistemically relevant to very many things, and therefore relevant to something each person values. Such are new discoveries about COVID-19: for each person, such discoveries probably shed light on something she or he values.

The relevant sense of aesthetic significance should be understood as a domain of fitting valuing, and one which is plausibly viewed as permissive. If both valuing opera and valuing thrash metal—two very different music genres—are justified, they are probably not required. Our account says that whether one values opera or thrash metal can matter to the fittingness of one's surprise. Watching Mozart's *Don Giovani* for the first time might be surprising to someone who values opera, but not for someone who, instead, values thrash metal. Similarly, listening to Dark Angel's "The Burning of Sodom" for the first time might fittingly surprise a thrash metal enthusiast but it would not fittingly surprise someone who would rather avoid metal music. In either case, one's fitting surprise is a response to a substantive aesthetic quality, but the response is merited partly due to the agent's broader aesthetic orientations, sensibilities, and commitments.

Comparative vs. Non-Comparative Significance. Suppose you buy most of the tickets in a thousand ticket lottery. Ticket number 822 is announced as the winner, and, lo and behold, it is one of the tickets you own. You have just won a large sum of money. Is surprise fitting? Intuitively the answer is 'no', surprise does not seem fitting. You should have expected to win given that you bought most of the tickets; it doesn't matter which particular ticket made you

win. However, it was unexpectable that ticket 822 will be drawn and the fact that it was drawn has a significant impact on what you value due to the sum of money you will receive. So our account seems to predict, implausibly, that the fact that ticket 822 was drawn is surprising.

However, our account avoids this result by employing a *comparative* notion of significance. The significance of ticket 822 is determined not merely by its impact on what the agent values, but by its impact on what the agent values *in comparison to the impact of alternative outcomes*. It is significant that one of the tickets you bought was drawn (relative to the alternative scenario in which none of the tickets you bought are drawn), but it is not significant that this particular ticket was drawn (relative to other tickets you bought). Since it was expectable that one of the tickets you bought will be drawn, the fact that one of them was drawn is not surprising.

Consider another example. The owners of a local restaurant decide to add to their set menu a rotating 'dish of the day'. Their menu announces: "Every day a new surprise!"[25] In this scenario, we know in advance that every day, including today, there will be an additional dish that is not on the set menu. It might be Indian curried tofu, Middle Eastern hummus with fava beans, Georgian Khachapuri or what have you. It is possible that all of these dishes are ones that you like, and as such they have some personal significance to you. Also, it may be that every day you fittingly do not expect the specific dish of the day. If significance is interpreted non-comparatively, then each dish of the day will be unexpectable and significant and therefore surprising. However, if significance is interpreted comparatively, and the significance of each of the possible dishes amounts to 'being a dish you enjoy', then it is not significant which particular dish is the dish of the day because you would have a dish you enjoy no matter which particular dish you are served. Therefore, none of the 'surprise' dishes are actually surprising.[26]

---

[25] Selim Berker presented us with an imaginary example with a similar structure. Then we came across this sign in a restaurant and decided to illustrate how our theory helps us make sense of the real world.

[26] Nevertheless, it is arguable that although each of the different 'surprise dishes' is a dish you like, each has different qualities and flavors and is therefore significant to you in a different way. The issue is not one of

V. IMPLICATIONS

A. Falk on coincidences

Our theory of surprisingness has implications for the interpretation of certain empirical findings. In a famous article, Ruma Falk (1989) demonstrates that people tend to be more surprised by coincidences involving themselves than by coincidences involving others. This is sometimes thought of as a bias, in the sense that it leads to unjustified judgements of surprise. For example, Diaconis and Mosteller (1989, 854) write that "Falk...showed that people found stories that happened to themselves more surprising than the same stories occurring to others...these findings agree with common sense...we all believe some things that are wrong"; and Johansen and Osman (2015, 37) interpret Falk as having discovered an "egocentric bias that influences the way people make probability estimates." These authors assume that the only factor that determines surprisingness is probability. Therefore, if probabilities are equal in two cases and people interpret one as being more surprising than the other, their judgment must be irrational.

A nice upshot of our significance account is that it offers an interpretation of Falk's findings that explains away the so-called egocentric bias. The phenomenon Falk describes need not be interpreted as irrational. On our account, it is generally fitting to be surprised by coincidences

---

incomparability—it need not be the case that no positive comparative judgment of the dishes' value is true (Chang 1997). Even if the different dishes are equally valuable they might be differently valuable. Compare two $20 bills to two equally delicious meals of different kinds. There is no relevant difference between the value of the $20 bills but there is a difference in the value of the delicious spaghetti compared to the delicious souvlaki. The difference might not be such as to give you a reason to prefer one to the other, but it might be significant nonetheless. Therefore, that the daily dish is curried tofu rather than hummus might be significant even if you would enjoy either one equally, for you might still enjoy them differently. If this line of reasoning is successful, and the differences between the daily dishes are significant, then today's 'surprise dish' might be surprising after all, even according to the comparative interpretation of significance.

that happen to you more than by coincidences that happen to others, even if probabilities are the same, because coincidences that happen to you tend to be more significant to you.[27]

B. Smith on *Rosencrantz and Guildenstern*

Martin Smith (2017), discussing a fictional example from *Rosencrantz and Guildenstern Are Dead* (Stoppard 1966), argues that if a fair coin were to land 92 times heads in a row, it would be irrational to be surprised. In such a case, he claims "we shouldn't feel surprised...we have no reason to feel surprised and, if we do feel surprised, then we're being irrational" (p. 2). We disagree. A good feature of our account is that it predicts that we should be surprised by such an occurrence. If a coin were to land heads so many times in a row, it would be both very unexpected and significant. It would therefore be surprising, according to our account.

Smith's argument relies on at least two premises that we reject. First, Smith claims that "it's rational to be surprised by an event if and only if that event requires investigation and explanation" (p. 7). Now, of course, if the coin tosses are known to be fair, and Smith seems to assume throughout that this is the case, then the only explanation of the sequence of heads can be mere coincidence and no further investigation is needed. In real life, we are never in a position to be fully sure of coin tosses that they are fair, but in imaginary cases anything can happen. It then would follow from Smith's premise that the sequence of 92 heads would not be surprising. Our theory of fitting-surprise makes no such prediction. To the contrary, it predicts that some facts are surprising even if they don't require investigation or explanation. We believe this is the correct result. If you win a million dollars in a million-ticket lottery, your surprise would be fitting independently of any need for investigation or explanation.

---

[27] Teigen & Keren (2003, 70) note in passing an interpretation of Falk's findings along these lines. However, their research is descriptive—it is concerned with what actually elicits judgments of surprise—not normative. They suggest that people are more surprised by coincidences that happen to them because they are more significant to them, but they are not concerned with whether this makes the response fitting.

Smith thinks that his first premise supports his crucial premise, which he calls the *conjunction principle*. According to this principle, "If it's unsurprising for event e1 to happen, and it's unsurprising for event e2 to happen, and these two events are independent of one another, then it's unsurprising for e1 and e2 to both happen" (p. 2). If the conjunction principle is true, then from the fact that for each individual toss it is unsurprising that it lands heads, it follows that the conjunction of tosses landing heads is unsurprising. Smith points out that the conjunction principle follows from some previous works on surprise, namely those of Shackle (1969, 80) and Spohn (2012).

In contrast, our theory of fitting-surprise predicts that the conjunction principle is false. We believe this is as it should be. The reason the conjunction principle is false is twofold. First, as is well known from probability theory, if e1 and e2 have high probability, it doesn't follow at all that their conjunction has high probability. The same is true of expectations, which in many ways act like probabilities. It can be true of every individual coin toss that I expect it to land heads just as much as I expect it to land tails. It doesn't at all follow that I expect a long sequence of tosses to land all-heads just as much as I expect the long sequence to not land all-heads. In a single toss, the coin can only land heads or tails. In a sequence of n tosses, there are 2^n possible sequences, only one of which is all-heads.

Another reason why the conjunction principle is false has to do with the second contributing aspect of what is surprising, namely, significance. It can be the case that e1 and e2 are each insignificant in themselves but in conjunction they are very significant. The fact that Elizabeth Bennet was at a particular ball at a particular time is not of any significance in itself. The fact that Mr Darcy was at a particular ball at a particular time is not of any significance in itself either. The conjunction turns out to be very significant. They would never have met and later fallen in love were it not for the fact that they both attended that ball. Similarly, it is true of any

individual toss that there is no significance to how the coin lands, whether heads or tails. However, it does not follow that there's no significance to a long sequence of heads.

Our conclusion is therefore the opposite of Smith's. If a fair coin were to land 92 times heads, this fact can indeed merit surprise.

C. Incomparability

We argued that fitting surprise is graded; some things merit more surprise than others. Winning 1,000,000€ in a thousand ticket lottery is less surprising than winning 1,000,000€ in a million-ticket lottery. These are two cases that are easy to compare. There are cases that are more difficult and perhaps impossible to compare. Is it more surprising to win 1,000,000€ in a thousand ticket lottery, or 1,000€ in a million-ticket lottery? We find it difficult to judge. Our account provides an explanation for this difficulty. Fitting surprise is determined by two factors, expectations and significance. It does not provide an answer to how the two factors weigh against each other. It is easy to compare two cases that differ only in expectation but not in significance. When the cases vary both in significance and in degree of expectation such that the factors pull in opposite directions, our account in its current form provides no guidance as to how they should be compared.

There can also be cases that vary in the type of significance at stake. Would it be more surprising to find out your best friend won 1,000,000€ in a large lottery or to see an elephant walking around campus? It is difficult to say. Our theory of surprise can explain why it might be difficult to answer this question even if the degrees of expectations are similar. It would be difficult to say because the kinds of significance involved in the two cases are very different.

Is comparison in these cases difficult or is it impossible? That is, is there a correct ordering of surprising facts in terms of how comparatively surprising they are, or are some cases deeply

incomparable? This is an interesting question which we leave open. We are open to there being incommensurable cases, which would imply that, while surprise is graded, it is not degreed and not properly representable by numbers.

But even if cases of fitting surprise are incommensurable they might still be comparable. As Ruth Chang has noted, while *incommensurable* items cannot be precisely measured against a single 'scale' of units of value, they might still be *comparable*, that is, one might be better or more choice-worthy than the other (Chang 2001).[28] Given the variety of kinds of significance, it is likely that surprising facts are often incommensurably surprising, but given that we can often compare the incommensurable significance of various facts, it is also likely that many incommensurably surprising facts are comparable.

VI.   CONCLUSION: THE MULTI-FACETED SIGNIFICANCE OF SURPRISE

To recap, we have argued that fitting surprise is not determined solely by unexpectability but also by significance. According to our view, surprise is merited when one discovers that the world is *significantly* not as one (fittingly) expected it to be. Therefore, we maintain that surprise should not be understood as an essentially or merely epistemic reaction.

Our proposal aligns with recent views in epistemology. Several epistemologists have suggested that knowledge and justified belief are "encroached" by moral or pragmatic considerations: whether a person knows or justifiably believes p is partly determined by non-epistemic considerations provided by the person's moral or pragmatic circumstances.[29] Without taking a stand on knowledge and justified belief, we propose that fitting surprise is partly determined

---

[28] "Comparable items can be ordinally ranked—ranked on a list—and need not be cardinally ranked—precisely ranked against some unit of value" (Chang 2001, 2).

[29] See, for example, Hawthorne and Stanely (2008), Fantl and McGrath (2009), Fritz (2017), Basu and Schroeder (2018), and Moss (2018).

by non-epistemic considerations provided by what the person fittingly values. Fitting surprise reflects more than our view of what *is*; it reflects our view of what *is significant*.

LIST OF WORKS CITED

Baldi, P. (2002). A computational theory of surprise. In *Information, Coding and Mathematics* (pp. 1–25). Springer US. doi:10.1007/978-1-4757-3585-7_1

Baras, D. (forthcoming). *Calling for Explanation*. Oxford University Press.

Basu, R., & Schroeder, M. (2019). Doxastic wrongings. In B. Kim & M. McGrath (Eds.), *Pragmatic encroachment in epistemology* (pp. 181–205). New York: Routledge.

Chang, R. (2001). Introduction. In *Incommensurability, Incomparability and Practical Reason* (pp. 1–34). Harvard University Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 181–204. doi:10.1017/S0140525X12000477

Clark, A. (2018). A nice surprise? Predictive processing and the active pursuit of novelty. *Phenomenology and the Cognitive Sciences*, 17(3), 521–534. doi:10.1007/s11097-017-9525-z

Currie, A. (2018). The argument from surprise. *Canadian Journal of Philosophy*, 48(5), 639–661. doi:10.1080/00455091.2017.1368860

D'Arms, J., & Jacobson, D. (2000a). Sentiment and value. *Ethics* 110(4): 722–748.Deonna, J., & Teroni, F. (2012). *The Emotions: A Philosophical Introduction*. Routledge.

D'Arms, J., & Jacobson, D. (2000b). The moralistic fallacy: on the 'appropriateness' of emotions. *Philosophy and Phenomenological Research* 61(1): 65–90. doi:10.2307/2653403

Diaconis, P., & Mosteller, F. (1989). Methods for studying coincidences. *Journal of the American Statistical*

Association, *84*(408), 853–861. doi:10.2307/2290058

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, *3*(4), 364–370. doi:10.1177/1754073911410740

Ely, J., Frankel, A., & Kamenica, E. (2015). Suspense and surprise. *Journal of Political Economy*, 123(1), 215–260. doi:10.1086/677350

Falk, R. (1989). Judgment of coincidences: Mine versus yours. *The American Journal of Psychology*, *102*(4), 477. doi:10.2307/1423303

Fantl, J., & McGrath, M. (2009). *Knowledge in an Uncertain World*. Oxford: Oxford University Press.

Frankfurt, H. (1982). The importance of what we care about. *Synthese*, *53*(2), 257–272. doi:10.1007/BF00484902

Friedman, J. (2013). Suspended judgment. *Philosophical Studies*, 162(2), 165–181. doi:10.1007/s11098-011-9753-y

Fritz, J. (2017). Pragmatic encroachment and moral encroachment. *Pacific Philosophical Quarterly*, 98(1), 643–661.

Gendolla, G. H. E. (1997). Surprise in the context of achievement: The role of outcome valence and importance. *Motivation and Emotion*, *21*(2), 165–193. doi:10.1023/A:1024486617134

Goldie, P. (2002). *The Emotions*. Oxford University Press. doi:10.1093/0199253048.001.0001

Harker, D. (2012). A surprise for Horwich (and some advocates of the fine-tuning argument (which does not include Horwich (as far as I know))). *Philosophical Studies*, *161*(2), 247–261. doi:10.1007/s11098-011-9732-3

Hawthorne, J., & Stanley, J. (2008). Knowledge and action. *The Journal of Philosophy*, 105(10), 571–590.

Horwich, P. (1982). *Probability and Evidence*. Cambridge University Press.

Howard, C. (2018). Fittingness. *Philosophy Compass* 13:e12542. doi:10.1111/phc3.12542

Howard, C. (2019). The fundamentality of fit. In *Oxford Studies in Metaethics Volume 14* (pp. 216–236). Oxford University Press. doi:10.1093/oso/9780198841449.003.0010

Jackson, E. (2020). The relationship between belief and credence. *Philosophy Compass*. doi:10.1111/phc3.12668

Johansen, M. K., & Osman, M. (2015). Coincidences: A fundamental consequence of rational cognition. *New Ideas in Psychology*, *39*, 34–44. doi:10.1016/j.newideapsych.2015.07.001

Kiesewetter, B. (2017). *The Normativity of Rationality*. Oxford University Press. doi:10.1093/oso/9780198754282.001.0001

Kolmogorov, A. N. (1950). *Foundations of the Theory of Probability*. Chelsea Publishing Company.

Lange, M. (2016). *Because Without Cause*. Oxford University Press. doi:10.1093/acprof:oso/9780190269487.001.0001

Lord, E. (2018). *The Importance of Being Rational*. Oxford University Press. doi:10.1093/oso/9780198815099.001.0001

Maguire, R., Maguire, P., & Keane, M. T. (2011). Making sense of surprise: An investigation of the factors influencing surprise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 176–186. doi:10.1037/a0021609

Manson, N. A., & Thrush, M. J. (2003). Fine-tuning, multiple universes, and the this universe objection. *Pacific Philosophical Quarterly*, *84*, 67–83. doi:10.1111/1468-0114.00163

Mellers, B., Fincher, K., Drummond, C., & Bigony, M. (2013). Surprise. A belief or an emotion? In *Progress in

Brain Research (1st ed., Vol. 202). Elsevier B.V. doi:10.1016/B978-0-444-62604-2.00001-0

Miller, C. R. (2015). *Surprise: The Poetics of the Unexpected from Milton to Austen*. Cornell University Press.

Morgan, M. S. (2005). Experiments versus models: New phenomena, inference and surprise. *Journal of Economic Methodology*, 12(2), 317–329. doi:10.1080/13501780500086313

Moss, S. (2018). Moral encroachment. *Proceedings of the Aristotelian Society*, 118(2), 177–205.

Munnich, E. L., Foster, M. I., & Keane, M. T. (2019). Editors' introduction and review: An appraisal of surprise: Tracing the threads that stitch it together. *Topics in Cognitive Science*, *11*(1), 37–49. doi:10.1111/tops.12402

Na'aman, O. (2021). The rationality of emotional change: toward a process view. *Noûs*, 55(2), 245–269. doi:10.1111/nous.12304

Parfit, D. (1998). Why anything? Why this? *London Review of Books*, 20(2).

Parfit, D. (2011). *On What Matters* (Vol. 1). Oxford University Press.

Scarantino, A., & de Sousa, R. (2021). Emotion. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Scheffler, S. (2010). *Equality and Tradition: Questions of Value in Moral and Political Theory*. Oxford University Press.

Schlesinger, G. N. (1991). *The Sweep of Probability*. University of Notre Dame Press.

Shackle, G. L. S. (1969). *Decision Order and Time in Human Affairs* (2nd ed.).

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x

Shogenji, T. (forthcoming). Probability and proximity in surprise. *Synthese*. doi:10.1007/s11229-020-02761-6

Smith, M. (2017). Why throwing 92 heads in a row is not surprising. *Philosopher's Imprint*, *17*(21), 1–8.

Spohn, W. (2012). *The Laws of Belief*. Oxford University Press.

Stoppard, T. (1966). *Rosencrantz and Guildenstern Are Dead*. Faber and Faber.

Teigen, K. H., & Keren, G. (2003). Surprises: Low probabilities or high contrasts? *Cognition*, *87*(2), 55–71. doi:10.1016/s0010-0277(02)00201-9

Tomkins, S. S. (1962). *Affect, Imagery, Consciousness*. Springer Publishing Co. doi:10.1037/14351-000

Weaver, W. (1948). Probability, rarity, interest, and surprise. *The Scientific Monthly*, 67(6), 390–392.

White, R. (2005). Explanation as a guide to induction. *Philosophers' Imprint*, 5(2), 1–29.

White, R. (2018). The argument from cosmological fine-tuning. In G. Rosen, A. Byrne, J. Cohen, S. V. Shiffrin, & E. Harman (Eds.), *The Norton Introduction to Philosophy* (2nd ed., pp. 29–35). Norton.