ORIGINAL RESEARCH

In defense of really statistical explanations

Marc Lange[1] (ORCID)

Abstract

According to Lange (2013, 2017), "Really Statistical (RS) explanations" constitute an important kind of non-causalscientific explanation. However, Roski (2021) has argued that all alleged RS explanations are either causalexplanations or not explanations at all. In so arguing, Roski has invoked Kahneman's (2011) interpretation of onealleged RS explanation. I employ Roski's arguments as an opportunity to elaborate and defend RS explanations. Iargue that "RS explanations" genuinely explain rather than deny the presuppositions of why-questions. I argue thatthe RS model is not excessively permissive in allowing some explanations to work purely statistically rather than bydescribing causal relations. I argue that Roski's view that some "RS explanations" operate by describing causalrelations fails to capture the kind of explanatory insight that RS explanations provide. I elaborate the notion of a"characteristically statistical phenomenon" that figures in the RS model and thereby explain how an RS explanationreveals that its explanandum is independent of any causal facts or specific chances, but rather depends only on ageneric sort of arrangement of chances. Finally, I argue that Roski misinterprets Kahneman, who actually holds thatthe explanation he discusses is a non-causal explanation that nicely fits the RS model. RS explanations constitutean important kind of non-causal scientific explanation.

Keywords Causation · Chance · Explanation · Regression toward the mean · Statistics

In previous work (Lange 2013, 2017), I have argued that "Really Statistical (RS) explanations" constitute an important kind of non-causal scientific explanation. How-

✉ Marc Lange
   mlange@email.unc.edu

1   Department of Philosophy, University of North Carolina at Chapel Hill,
    27599-3125 Chapel Hill, NC, USA

 Springer

ever, Roski (2021) has responded to Taylor's (2018) appeal to these explanations by denying that there are RS explanations. Rather (Roski has argued), all alleged RS explanations are either causal explanations or not explanations at all. In so arguing, Roski has invoked Kahneman's (2011) interpretation of one alleged RS explanation.

Roski's arguments provide an opportunity for me to elaborate and defend the RS model. I will argue that Roski has failed to show that various alleged RS explanations are either causal explanations or non-explanations. I will also argue that Kahneman's view of the alleged RS explanation he discusses is that (contrary to Roski) it is indeed a non-causal explanation that fits the RS model nicely.

An RS explanation (according to Lange 2013, 2017) derives its explanatory power by revealing that the fact to be explained is an instance of some particular, characteristically statistical phenomenon (such as regression toward the mean). Although (in Lange 2013, 2017) I gave some examples of characteristically statistical phenomena, I gave no general specification of what such a phenomenon is. Below I will give two necessary conditions for being a "characteristically statistical phenomenon". I will use these conditions to supplement my response to Roski by identifying the kind of explanatory insight that RS explanations supply in science and that cannot be supplied by descriptions of causal histories, no matter how ideally complete they may be.

In § 1, I will review the account of RS explanations that I have given previously (Lange 2013, 2017) and provide some new examples of putative RS explanations to consider. In § 2, I will argue that contrary to Roski, some "RS explanations" genuinely explain rather than deny the presuppositions of why-questions. In § 3, I will argue that the RS model is not excessively permissive in allowing some scientific explanations to work purely statistically rather than by describing causal relations. I will also argue that an interpretation of some "RS explanations" as working by describing causal relations fails to capture the kind of explanatory insight that RS explanations provide. Crucial to that insight is the way that an RS explanation identifies a particular type of characteristically statistical phenomenon (such as regression toward the mean) and reveals the explanandum to be an instance of that phenomenon. In § 4, I will give a new account of what a "characteristically statistical phenomenon" is. An RS explanation shows the explanandum to be likely given some generic sort of statistical arrangement and independent of the particular values of the chances involved, the causal relations, and the causal laws. Finally, in § 5, I will argue that it is really I, rather than Roski, who is entitled to appeal to Kahneman for support. Kahneman should be understood as arguing that we can best fit scientific practice by recognizing RS explanations as an important kind of *non-causal* scientific explanation.

1 RS explanations

As I argued in Lange (2013, 2017), some RS explanations appeal to regression toward the mean. In Galton's original example of regression toward the mean, exceptionally tall parents tend to have children who (although taller than average) are not as exceptionally tall. The RS explanation is that although parental heights and children's heights are correlated, the correlation is imperfect; an individual's height reflects not only inherited genes, but also some degree of chance. For a parent to reach an excep-

tional height, many chance processes must each end up with the outcome that most contributes to height. This is less likely than that some of these processes do whereas others do not end up with the outcome that most contributes to height. Therefore, the offspring of parents who are exceptionally tall (or exceptionally short) are apt to have heights that are less extreme – i.e., that "regress toward the mean."

Regression toward the mean is likely to happen in any case where two types of outcomes are imperfectly correlated. I gave the following example in Lange (2017:191):

> [S]uppose a fair coin is tossed 100,000 times. Consider the various runs of 20 consecutive tosses beginning with toss numbers 1, 11, 21,…; neighboring runs share 10 tosses, so a run with more than 10 heads tends to be followed by another run with more than 10 heads. Nevertheless, a run with an exceptionally high number of heads (let's say, 18 or above) tends to be followed by a run with fewer heads. This result is explained by regression toward the mean.

My point was that although this result could also be given a causal explanation (i.e., "is also explained by the coin's 50% chance of landing heads on any given toss (independently of the outcomes of other tosses)"), that causal explanation is not "Really Statistical" (even though it involves chances) since it works by taking the coin's 50% chance of landing heads and computing from it the chance that a run with an exceptional number of heads will be followed by a less extreme run. This explanation is causal because it acquires its explanatory power by virtue of "describing relevant features of the result's causal history." By contrast, the explanation by regression toward the mean "depicts the result as fallout from the statistical character of the case: not from the 50% chance of a toss's landing heads, not even from the chances of a 10-toss run's having various numbers of heads, but rather from the mere fact that there is a statistical association between the outcomes of overlapping 20-toss runs … The point of the explanation is [not to describe the result's causal history but] instead to exhibit the [explanandum] as arising from the fact that successive [overlapping 20-toss] runs" are imperfectly correlated, regardless of the precise "relation or its (perhaps probabilistic) causes – or, indeed, whether it has any causes at all" (Lange 2017:192).

I argued in Lange (2013, 2017:190-6) that we can explain the fact that most children of exceptionally tall parents are not so exceptionally tall by identifying this result as an instance of regression toward the mean – that is, as an instance of this particular "characteristically statistical phenomenon." This explanation is "non-causal" in the sense that it does not acquire its explanatory power by virtue of providing information about causal relations. Instead the explanation works by revealing that the explanandum is just (in the words of one textbook) "a statistical fact of life" (Gravetter and Wallnau 2009:536; cited by Lange 2017:190) and by identifying which particular "statistical fact of life" it instantiates. For this reason, I dubbed it a "Really Statistical (RS)" explanation. I said that "an explanation is RS if and only if it works by identifying the explanandum as an instance of some characteristically statistical phenomenon" (Lange 2017:196). I will be discussing all of this further in subsequent sections.

There are many other "characteristically statistical phenomena" besides regression toward the mean. Another one that I mentioned (Lange (2017)) as figuring in some RS explanations is that when chances govern various outcomes, small samples have a greater tendency than large samples to depart greatly from the expected value of some quantity; Kahneman (2011) refers to this fact as "the Law of Small Numbers." I cited (Lange 2017:193) Rutherford and Geiger as having used this "statistical fact of life" to explain why it is that "in counting the α particles emitted from radioactive substances … the average number of particles [emitted] from a steady source is nearly constant, when a large number is counted, [but] the number appearing in a given short interval is subject to wide fluctuations" (Rutherford and Geiger 1910:698). As we will see, Kahneman cites the Law of Small Numbers as explaining why the US counties with the lowest and highest incidences of kidney cancer are mostly (and disproportionately) counties with relatively small populations.

Unsurprisingly, Darwin gave many RS explanations. For instance, Darwin considered why wealthy breeders (of horses, dogs, sheep, and so forth) tend to be more successful than less wealthy breeders in breeding useful varieties. Darwin's explanation is that wealthy breeders tend to have larger stocks and (here comes the RS explanation) it is "just statistics" that the chance that a useful variety will arise is greater insofar as the population of animals being bred is larger. Darwin (1872:29) wrote: "as variations manifestly useful or pleasing to man appear only occasionally, the chance of their appearance will be much increased by a larger number of individuals being kept." Likewise (Darwin 1883:221), "Lord Rivers, when asked how he succeeded in always having first-rate greyhounds, answered, 'I breed many, and hang many.'"[1]

If RS explanation is a genuine type of scientific explanation, then it is a scientifically important kind of explanation. So it is philosophically important to understand the way in which alleged RS explanations operate. This is what Roski investigates.

2 "RS explanations" that are not genuine scientific explanations?

Regarding the examples that I have interpreted as RS explanations (and hence as non-causal explanations), Roski (2021:14133) argues that they "are either no genuine explanations or else convey information about causes" and should therefore be understood as causal explanations.

Consider, for instance, the putative explanation of the fact that the counties with the lowest and highest incidences of kidney cancer are mostly (and disproportionately) counties with relatively small populations. The purported explanans is the "Law of Small Numbers". Roski, however, says that this "purported explanation is not genuine" (p. 14133). Roski admits that his view needs to explain away the fact that this non-explanation appears to be an explanation; his view needs to "account for the fact that the information provided by [the Law of Small Numbers] seems to resolve a certain type of puzzlement – indeed, a puzzlement that often triggers a why-question" (p. 14133), namely, "Why is there this 'curious correlation between population … and cancer incidence'?" (p. 14133). According to Roski, the Law of

---

[1] For more on Darwin's explanation, see Beatty 2006.

Small Numbers does not answer this why-question by providing an explanation. Rather, the answer involving the Law of Small Numbers denies a presupposition of the why-question. This presupposition is "that there is some sort of influence between population density and cancer incidence" (p. 14133) – a causal influence. The answer invoking the Law of Small Numbers denies that there is any such influence: "The law answers the why-question, but not by providing an explanation … [T]he law removes a presupposition of the initial request, namely, that there is some relation of influence between population density and cancer incidence" (p. 14133).

I recognize that responding to a why-question by providing an explanation is distinct from responding by denying one of the question's presuppositions. For instance, the question "Why is p the case?" presupposes that p is indeed the case, and so the response "In fact, it is not the case that p" denies a presupposition of the question without providing an explanation of p. I therefore agree with Roski that "removing a puzzlement that triggered a why-question is not, by itself, providing an explanation" (p 14134). I also agree that someone asking why the counties with the highest and lowest frequencies of kidney cancer are mostly (and disproportionately) counties with relatively low populations may well be expecting the answer to reveal some remarkable causal connection between population and cancer.

However, I reject Roski's argument that in this example, the Law of Small Numbers is denying a presupposition of the why-question rather than supplying a genuine explanation of the observed association. Roski may be correct that in some contexts, this why-question carries the tacit presupposition that there is some sort of causal connection between population and cancer frequency. But this presupposition (if it exists) can easily be cancelled (i.e., eliminated) by being made explicit and then being disavowed, as in the following:

> I do not want to assume that there is some causal connection between population and cancer frequency. But we have seen that extreme cancer frequencies (high and low) are mostly (and disproportionately) present in counties with relatively low populations. Why is there this association?

This why-question (in the context of the remarks preceding it) fails to presuppose a causal connection between county population and kidney cancer rate. The correct answer to this why-question is that the Law of Small Numbers explains the observed association. There is no reason to regard this answer as rejecting a presupposition of the why-question rather than as supplying an explanation of the observed association.

(Obviously, we cannot use the same sort of disavowal to eliminate from the question "Why is p the case?" the presupposition that p is the case. It is pragmatically infelicitous to say "I do not want to assume that p is the case. But why is p the case?" By contrast, there is no such pragmatic infelicity when we ask the why-question about population and cancer frequency after cancelling any presupposition that these factors are causally connected.)

Undoubtedly, there are why-questions correctly answered by RS explanations where the questioners initially expected some sort of causal explanation instead. When a teacher wonders why the students who performed best on the course's first exam tended not to be the students who performed best on its second exam, the

teacher may well initially have expected to uncover some causal explanation, such as that the best-performing students on the first exam became overconfident and so did not study hard for the second exam. But suppose that the teacher notices that in addition, the students who performed *worst* on the course's first exam tended not to be the students who performed worst on the second exam. This gives the teacher some reason to doubt that a causal explanation would give the whole story regarding the first-exam's *best*-performing students' tending not to perform best on the second exam. (Overconfidence going into the second exam surely did not plague the first exam's *worst* performers.) The teacher then asks why the first exam's best-performers tended to do a little worse on the second exam *and* the first exam's worst performers tended to do a little better. The teacher couples these two results because the combination is suspicious; she suspects that they may have some sort of *common* explanation. Yet (contrary to Roski's suggestion) she also *doubts* the existence of any common *causal* factor. (Again, overconfidence going into the second exam cannot plausibly have been a common causal factor afflicting both the first exam's best performers and its worst performers.) The result regarding the first exam's worst performers makes more puzzling the result regarding the first exam's best performers precisely because there is presumably no explanatory common *cause* for both student groups. Regression toward the mean (an RS explanation) removes the puzzlement by providing the suspected explanation common to both student groups without identifying an explanatory common *cause*. This is a case of an RS explanation where (contrary to Roski) the why-question (concerning both student groups) does *not* presuppose the existence of causal connection between being one of most extreme performers on the first exam and not so being on the second exam.

3 "RS explanations" that are causal explanations?

Having rejected Roski's argument that some alleged RS explanations are not genuine explanations (but instead merely remove a presupposition of the why-question), I will now turn to the examples that I have interpreted as RS explanations but that Roski argues are in fact causal explanations. These include cases that I have regarded as non-causal explanations invoking regression toward the mean, such as in the example involving runs of 20 consecutive coin tosses. Consider some statistical principle (such as that whenever there is an imperfect correlation between two results of chance setups, regression toward the mean is likely to occur). Roski maintains that if such a principle "is considered relevant for the explanation of a pattern manifested by the outcome of a series of throws, this is because it conveys information about the causal powers of the [system] in question, albeit highly unspecific information" (p. 14135), rather than because it supplies a non-causal explanation. In support of this view, Roski points out that there are propositions about the causal powers operating in the coin tosses (to use that example) that would, if true, preclude that result from being explained by regression toward the mean.

   I recognize that there are possible causal considerations that would preclude such an explanation. For example, suppose that (as Roski imagines (p. 14135) regarding a similar case) the coin tosses "were in fact not independent but rather rigged." Let's

say (embroidering Roski's example) that a powerful wizard deterministically caused the outcome of each coin toss; by making certain 20-toss runs have fewer heads than others, he made the outcomes appear as if regression toward the mean were occurring. (The wizard could just as easily have made each run contain exactly the same high number of heads as every other run, in apparent defiance of regression toward the mean.) In this sort of case, as Roski points out, regression toward the mean would not explain the outcome, even though the outcome is as if regression toward the mean had occurred. Roski's point is that even in an ordinary case without a wizard, an appeal to regression toward the mean contains some (albeit limited) information about the causal powers at work (e.g., the information that no such wizard deterministically caused the outcomes) and so the explanation supplied is causal.

But the fact that an explanation supplies some information about causes (and so would be precluded by some contrary causal information) does not automatically make the explanation "causal", since the explanation may not be deriving its explanatory power (even partly) by virtue of supplying information about the causes at work.[2] RS explanations derive their explanatory power from revealing that the fact

---

[2] I have made this point (in Lange 2017:19–20) in connection with a non-causal explanation that is not RS. (I have also made this point (in Lange 2017:192) in connection with RS explanations.) Consider the fact that Mother fails to divide her strawberries evenly among her children without cutting any. This fact can be explained by the fact that she had 23 strawberries and 3 children and that 3 does not go evenly into 23. This non-causal, "distinctively mathematical" explanation happens to cite causes of Mother's failure (the numbers of strawberries and children) and rules out various hypotheses about its causes. But that is not enough to make it a causal explanation because it derives its explanatory power other than by supplying information about the world's network of causal relations. (I have argued for this use of "causal explanation" by contrast with some philosophers' broader use according to which any explanation that provides information about the explanandum's causes automatically qualifies as a causal explanation.) A non-causal explanation (such as this one) can supply information about causes *incidentally*, not as (part of) the way it acquires explanatory power. It qualifies as an explanation (i.e., it derives its explanatory power, the explanation "works") by virtue of fulfilling a sufficient condition for being explanatory (a "model" of scientific explanation) that does not expressly require information about causes.(Not every set of sufficient conditions for being explanatory counts as a (correct) "model" of scientific explanation, though every model should provide sufficient conditions for being explanatory. A "model" of scientific explanation corresponds to a natural kind of explanation. It specifies what it is in virtue of which various explanations explain. We could generate a set of sufficient conditions by, for example, disjoining the sufficient conditions specified in two, wildly dissimilar models or by adding an arbitrary condition to the sufficient conditions specified in a model. But the result would not be a model of scientific explanation.)That an explanation cites causes is not *sufficient* to make it causal; its citing causes is also not *necessary* in order for it to be causal. As I have argued (Lange 2017:17), "when we explain why some body is moving uniformly (rather than nonuniformly) by noting that the body is experiencing no forces, we are not giving the explanandum's causes (since it has none). But we are explaining by virtue of describing a relevant aspect of the world's network of causal relations. … That there are no forces acting on the body qualifies as explanatorily relevant by virtue of the fact that forces cause accelerations. … We have here a causal explanation because the facts that explain are explanatorily relevant by virtue of their significance regarding the world's network of causal relations."Regression toward the mean might seem like uniform motion (as a referee suggested) in that both are the "default" behavior, i.e., the result that will (likely) occur in the absence of disturbing factors. But the initial departure from the mean (which is followed by regression toward the mean) does not occur because of a "disturbing factor" that later ceases, allowing regression to occur – unlike accelerated motion, which does occur because of a disturbing factor (a force), the departure of which would allow a return to unaccelerated motion. Regression toward the mean is the likely result of a certain general arrangement of chances (namely, the imperfect correlation of two outcomes), whatever the causal network might be like, whereas unaccelerated motion is explained by a feature of the causal network (namely, that accelerations require causes, which are forces). When a short run of tosses of a fair

being explained is just a "statistical fact of life" in being an instance of some characteristically statistical phenomenon such as regression toward the mean. (I will examine the idea of a "characteristically statistical phenomenon" in the following section.) That regression toward the mean cannot be occurring if certain causal relations obtain (e.g., if a wizard is deterministically causing the coin-toss outcomes) does not mean that an appeal to regression toward the mean derives its explanatory power by virtue of conveying that no such causal relations obtain. To interpret an RS explanation as deriving its explanatory power from the causal information it supplies would be to neglect the most explanatorily significant information supplied by an appeal to regression toward the mean, namely, the particular "characteristically statistical phenomenon" ("statistical fact of life") that the explanandum instantiates.[3] (Regression toward the mean is obviously not the only "statistical fact of life" that would be precluded by a wizard's intervention; a wizard who determined each of the cases of kidney cancer in each county could cause a county-by-county pattern that appears to reflect the Law of Small Numbers but in fact has no RS explanation.)

In characterizing the coin-toss explanandum as an instance of regression toward the mean, the RS explanation attributes the coin-toss explanandum to the same principle that is responsible for various other outcomes, including that the very tallest parents tend not to have the very tallest children, that the students who performed the very worst on the course's first exam tend not to be the students who performed the very worst on the second exam, and that the athletes whose extraordinary achievement during one season led to their appearance on the cover of *Sports Illustrated* usually appear to suffer a "jinx" by being less successful during the following season. Physically, these facts are obviously very diverse. The causes underlying any one of them have virtually nothing to do with the causes at work in any other. In explaining these cases as all instances of regression toward the mean, we reveal that at a deep level, they all have the same sort of explanation.[4] These explanations cannot be revealing an important underlying similarity among these cases if each of these various appeals to regression toward the mean explains merely (as Roski says) by conveying some information about the causes at work in that particular case: the particular causes operating in one of these cases are quite different from those operating

---

coin contains a high fraction of heads, but then as the run lengthens, the overall fraction of heads tends to be nearer to 50%, there has been no change in the causes between the initial short run and the later, longer run – whereas, by contrast, when an accelerating body later stops accelerating, the causes have changed.

[3] "[A]n RS explanation says a good deal more about the explanandum than merely that 'It's just statistics.' An RS explanation identifies the explanandum as an instance of some particular kind of behavior that is characteristic of statistical systems in virtue of their being statistical" (Lange 2017:194).

[4] I am not suggesting that an RS explanation's explanatory power derives from the unification it supplies (as claimed by unificationist accounts of scientific explanation). I am suggesting only that RS explanations supply a kind of unification in revealing that various (physically diverse) phenomena have the same sort of explanation. I recognize that (as a reviewer pointed out) two episodes of (e.g.) inflation may be physically very diverse at a lower level (e.g., because the currency being used is silver in one case and electronic in the other) and yet the two inflationary episodes may have the same sort of *causal* explanation in economics. However, two instances of regression toward the mean may have nothing in common even in terms of the natural kinds and properties of a *higher-level* scientific field, other than that each episode involves an imperfect correlation between two types of outcomes.

in the others. (The absence of wizards in each case is not an important underlying similarity.) In the next section, I will return to this causal diversity.

Although an RS explanation derives its explanatory power from identifying the particular "statistical fact of life" instantiated by the explanandum, Roski fails to note this key feature of RS explanations. Instead, he says that according to those philosophers who regard RS explanations as genuinely explanatory, an RS explanation explains simply by showing the explanandum to be likely, given the underlying chances. Accordingly, Roski objects that (unless it explains by virtue of the information it supplies about causes) an RS "explanation" is not genuinely explanatory: "Explaining why a particular distribution … prevails requires more than pointing to the mere fact that such a distribution is likely, statistically speaking" (p. 14133). But (I have just emphasized) the most explanatorily significant information provided by an RS explanation is the characteristically statistical phenomenon (e.g., regression toward the mean, the Law of Small Numbers) that the explanandum instantiates.

Because Roski neglects this key feature of RS explanations, he regards the RS model as vulnerable to the same fatal objection as the statistical relevance (SR) model of explanation encountered. Roski correctly says that "[t]heories of explanation … that take mere statistical relevance as sufficient for explanatory relevance are now widely dismissed" (p. 14134); the fact that $ch(e|c)=n>m=ch(e|{\sim}c)$, together with other facts about the statistical relevance to e of various circumstances (along with other statistical relations, such as that e cannot be screened off from c by other sorts of factors), are not enough to make c explanatorily relevant to e. Rather, Roski says, the causal relations frequently lying behind these statistical-relevance relations are indispensable to explanatory power (pp. 14134-5). I agree with Roski (and the general philosophical consensus) that the right lesson to draw from the failure to reduce causal relations entirely to statistical relations is that the SR model fails; the statistical-relevance relations that the SR model deems to be sufficient for scientific explanation can fail to correspond to causal relations, and when that happens, scientific explanation is associated with the causal relations rather than the statistical relations.

But none of this suggests that the RS model suffers from the same defect as the SR model, since the RS model does not aim to use statistical relations to underwrite *causal* explanations. The point of an RS explanation is to explain precisely *not* by fitting the explanandum into the world's causal network, but rather (as I am about to discuss) by revealing that the explanandum would (likely) result no matter what the causal network is like (or even if there is none) as long as there exists the relevant arrangement of chances. For example, as long as the heights of parents and offspring are imperfectly correlated, regression toward the mean will likely occur whether parental heights help to cause offspring heights, or vice versa, or neither.

4 What is a "characteristically statistical phenomenon"?

My official formulation in (Lange 2013, 2017) is that "an explanation is RS if and only if it works by identifying the explanandum as an instance of some characteristically statistical phenomenon" (Lange 2017:196). There I gave several examples of such phenomena, including

- regression toward the mean, which is likely if two types of outcomes are imperfectly correlated (p. 191);
- differences among equal-length runs in the frequency of a given outcome, which is likely if the runs are short and consist of repeated identical, independent trials (p. 193);
- small samples departing more often than large samples from the average frequency of some outcome, which is likely if the samples consist of repeated identical, independent trials (p. 193);
- 20-step walks on the integer number line (each walk starting at zero and moving one unit per step) where a majority of the walks remain on the same side of zero for at least 19 steps, which is likely in a large collection of random walks with a fixed chance of each step's direction (p. 195).

In Sect. 1, I noted another (invoked by Darwin): that a rare trial result occurs at least once more frequently during a large number of identical, independent trials than during a small number.

But in previous work, I said little about what all of these examples have in common that makes each of them a "characteristically statistical phenomenon". My other previous slogans -- that an RS explanation "depicts the result as fallout from the statistical character of the case" (p. 192) and "identif[ies] some particular signature of statistical phenomena that the explanandum exemplifies" (p. 195) – likewise leave unclear what counts as "the statistical character of the case" or a "signature of statistical phenomena". It would be better to have a general characterization of what it is to be a "characteristically statistical phenomenon."

Each entry in the above list is a *type* of characteristically statistical phenomenon. Each entry takes the form "result R, which is likely if there is some arrangement A of chances." An RS explanation characterizes the explanandum as an instance of one such R and as having resulted from the corresponding A. This is what it is to identify the explanandum as an instance of some characteristically statistical phenomenon.

One necessary condition for a "characteristically statistical phenomenon" is that the result R must be likely whatever the particular values of the chances in the given arrangement A. In other words, A does not specify the particular value of any chance -- or that its value falls within a given range, that its value is greater than some minimum, that one possible outcome's chance is greater than another's, etc. The arrangement of chances is *generic*: it specifies merely that there are (non-extremal) chances of (or that there are statistical dependencies between) various outcomes. For instance, some A might specify only that there is an imperfect correlation between two types of outcomes or that there are repeated identical, independent trials. That the explanandum is made likely given merely such a generic arrangement of chances – that the

explanandum is independent of (for instance) the particular values of various chances – is part of what an RS explanation reveals.

This necessary condition for a "characteristically statistical phenomenon" prevents RS explanations from encompassing the sort of non-explanations that concern Roski (as we saw at the close of the previous section) – i.e., non-explanations that mistakenly qualify as explanatory under the SR model. That $ch(e|c) = n > m = ch(e|{\sim}c)$ and other facts about the statistical relevance to e of various circumstances (or even simply that e is likely given c and unlikely given ${\sim}c$) is the sort of specific information about the chances that A cannot contain.

Because this necessary condition limits the information about chances that the arrangement A in an RS explanation can include, there are some outcomes of chance setups that cannot be given RS explanations. For instance, in the case given earlier of a fair coin and its successive, overlapping, 20-toss runs, we can use regression toward the mean to explain why runs with exceptionally high numbers of heads tend to be followed by runs with fewer heads. But we cannot use regression toward the mean (or any other RS explanation) to explain why a given 20-toss run (which starts with the final 10 tosses of a 20-toss run having an exceptionally high number of heads) has the particular frequency of heads it does or to explain any particular outcome in that run (e.g., to explain why the 12th coin toss landed "heads").

A second necessary condition for a "characteristically statistical phenomenon" is that A cannot specify any causal relations or natural laws. The generic statistical arrangement alone suffices to make the explanandum likely. No laws of nature, but rather the laws of probability alone, make the result R likely given A. For instance, when Rutherford and Geiger (1910:704) discovered that departures from a radioactive sample's mean rate of decay tend to be smaller over larger intervals, they said that this result is explained solely by "the laws of probability and that the α particles are emitted at random".

This second necessary condition elaborates my previous characterization of RS explanations as explaining by revealing the explanandum to be a "statistical fact of life." Because the information admissible into A is so limited, an RS explanation can be surprising in that it can seem something of a wonder that any result R as substantial as the explanandum can be rendered likely by an arrangement as meager as A. Furthermore, since the generic arrangement A cannot specify causal relations, an RS explanation is compatible with great heterogeneity among the causes at work in the various individual events that constitute R. Consider again the instance of regression toward the mean where the best-performing students on the first exam tend not to perform exceptionally well on the second exam. For a particular student, there are causes of her doing less than exceptionally well on the second exam; for instance, perhaps she slept well the night before the first exam but did not sleep well the night before the second exam. The causes of this student's doing less than exceptionally well on the second exam will typically be distinct from and dissimilar to the causes of another of the first exam's top performers doing less than exceptionally well on the second exam (such as that she was overconfident, ill, hungry from failing to eat breakfast, distracted by personal problems, or etc.). When we take all of the top-performers on the first exam, the causes of their generally performing less exceptionally on the second exam form a heterogeneous group of circumstances, differing from student to

student. This is often the case when we consider the causes of the individual events that constitute an instance of the result R in a "characteristically statistical phenomenon". The second necessary condition permits this causal heterogeneity.

The second necessary condition requires that A make R likely solely by virtue of the laws (i.e., the mathematics) of probability, not the laws of nature. The laws of probability possess a stronger sort of necessity than the laws of nature do; the laws of probability would still have held, even if (for instance) gravity had declined with the inverse-cube rather than with the inverse-square of the distance. Thus, in an RS explanation, that A makes R likely "transcends" the causal details of the case: no matter what those causal details might have been, R would still have been likely given A. In this way, an RS explanation rises above whatever "host of petty independent influences" (Hacking 1990:185) -- and even whatever causal laws -- may help to causally explain a particular event that partly constitutes a given R (such as, for instance, when a die's landing on six is one outcome contributing to R and is caused partly by the host of petty independent influences consisting of the die's various collisions with individual air molecules). In revealing the explanandum's independence from these causal details and its dependence solely on the generic statistical arrangement A, an RS explanation expresses a characteristically statistical phenomenon's "autonomy" (Hacking 1990:180−88) from whatever underlying causal to-ing and fro-ing there may be. By appealing solely to a generic statistical arrangement A, an RS explanation is "in some way … not reducible to some set of underlying causes" (Hacking 1990:181). An RS explanation thereby provides a kind of explanation that could not be supplied by even an ideally complete description of the causal histories of the events constituting R.

We should accept the RS model because insisting that the facts purportedly explained by RS explanations have only *causal* explanations not only runs contrary to scientific practice, but also fails to capture the distinctive kind of explanatory insight that RS explanations supply in science. Roski denies this, as I will now discuss further.

5 Kahneman on RS explanations as non-causal explanations

Roski appeals to Nobel-laureate Daniel Kahneman's (2011:109–112) remarks concerning one alleged RS explanation as evidence that science regards RS "explanations" as not genuinely explanatory. However, Roski misunderstands Kahneman's view. In fact, Kahneman's point is precisely that the given alleged RS explanation is a non-causal "statistical" explanation.

Kahneman's explanandum is the (now familiar) fact that the US counties where the frequency of kidney cancer is lowest or highest are mostly (and disproportionately) relatively unpopulated. Recall that on Roski's interpretation, the appeal to the Law of Small Numbers is no explanation of this fact, but instead denies a presupposition of the why-question. Roski (p. 14134) says that

> [t]his interpretation of the case is in line with how Kahnemann [sic] … conceives of the case:

The deeper truth is that there is nothing to explain. The incidence of cancer is not truly higher[5] than normal in a county with a small population. … [T]he differences between dense and rural counties do not really count as facts: they are what scientists call artifacts, observations that are produced entirely by some aspect of the method of research – in this case, by differences in the sample size. (Kahneman 2011:111)

.Roski (p. 14134) says that "Kahnemann [sic] dismisses the purported explanation of the relation between population density and cancer incidence in terms of the Law of Small Numbers."

But Kahneman does nothing of the kind. He regards the purported explanation not only as a genuine explanation, but also as a non-causal explanation – precisely contrary to Roski's view that alleged RS explanations are either non-explanations or causal explanations. With regard to the Law of Small Numbers explaining the association of extremely high and low kidney-cancer rates with low populations, Kahneman (p. 111) endorses this explanation and says the following about it (in a passage that Roski does not quote, though it appears on the same page as the passage he quoted above):

The explanation I offered is statistical: extreme outcomes (both high and low) are more likely to be found in small than in large samples. This explanation is not causal. The small population of a county neither causes nor prevents cancer…

Plainly, Roski is incorrect in characterizing Kahneman as "dismissing the purported explanation." Rather, Kahneman embraces it as a genuine explanation. It is I, rather than Roski, who is entitled to cite Kahneman as agreeing with *him*; Kahneman's view is that this is a non-causal explanation. This is Kahneman's (p. 110) main point:

the main lesson to be learned is not about epidemiology, it is about the difficult relationship between our mind and statistics. System 1 is highly adept in one form of thinking – it automatically and effortlessly identifies causal connections between events, sometimes even when the connection is spurious. When told about the high-incidence counties, you immediately assumed that these counties are different from other counties for a reason, that there must be a cause that explains this difference. As we shall see, however, System 1 is inept when faced with "merely statistical" facts, which change the probability of outcomes but do not cause them to happen.

Clearly, Kahneman takes one such "merely statistical" fact to be the small populations of certain counties, which raise the chance that their kidney-cancer frequencies will be extreme. (Kahneman's talk of "merely statistical" facts -- and of his explanation as "statistical" -- is very like my terminology of "really statistical" explanations and "statistical facts of life.")

---

5 The Kahneman text actually says "… truly lower or higher…"; Roski slightly misquotes it.

What about Kahneman's remark (quoted by Roski) that "there is nothing to explain. The incidence of cancer is not truly lower or higher than normal in a county with a small population…" (p. 111)? Kahneman's point, I believe, is *not* to dismiss the purported explanandum (that the counties with extreme frequencies of kidney cancer are mostly (and disproportionately) relatively unpopulated) as no fact at all (and thereby to dismiss its purported RS explanation). Rather, Kahneman is distinguishing this explanandum (which concerns *frequencies* and has an RS explanation) from the non-fact that counties with extreme frequencies of kidney cancer are (mostly) counties where an individual's *chance* of getting kidney cancer is extreme. Kahneman's point is that the extreme kidney-cancer *frequencies* in these counties are not due to the individuals in these counties having especially high or low *chances* of getting kidney cancer. (Their chances of getting kidney cancer are no different from anyone else's anywhere else.) Instead the extreme kidney-cancer frequencies in these counties are explained by the low "sample sizes" (i.e., populations) in these counties. (Unfortunately, Kahneman uses the term "the incidence of cancer" ambiguously; sometimes he means the *frequency* (i.e., the rate, the fraction) of individuals with cancer, but on other occasions, he means the objective *chance* of an individual's getting cancer.)

That this is Kahneman's point is somewhat clearer when we look at the complete Kahneman passage, including the part that Roski replaced with an ellipsis:

> The deeper truth is that there is nothing to explain. The incidence [that is, the chance – ML] of cancer is not truly lower or higher than normal in a county with a small population, it just appears to be so in a particular year because of an accident of sampling. If we repeat the analysis next year, we will observe the same general pattern of extreme results in the small samples, but the counties where cancer was common last year will not necessarily have a high incidence [that is, a high frequency – ML] this year. If this is the case, the differences between dense and rural counties do not really count as facts: they are what scientists call artifacts… (p. 111)

That *different* small-population counties have extreme cancer *rates* in different years strongly suggests that the counties with extreme cancer rates in a given year are not counties where an individual has an extreme *chance* of getting cancer.

When Kahneman says that "there is nothing to explain," he means that there is no especially high (or low) *chance* of cancer in a given individual living in the counties with especially high (or low) cancer frequencies in a given year. (Such an association of extreme cancer chances with low county populations would be in need of an explanation.) The appearance (from the extreme cancer frequencies) of extreme cancer chances in certain counties is indeed misleading – an artifact of the small populations in those counties. But Kahneman does *not* mean that there is literally nothing to explain; there remains the association of extreme county cancer rate with low county population. It has no causal explanation, according to Kahneman, but it has an RS explanation. In particular, it is "produced entirely by … differences in sample size" (p. 111). Kahneman is emphatic that the stable pattern of extreme results in low-population counties has (contrary to Roski) a non-causal explanation.

6 Conclusion

I have elaborated the RS model and thereby argued that Roski has found no reason to suspect that "RS explanations" are either causal explanations or not explanatory at all. I have tried to clarify how RS explanations work: the way that they acquire their explanatory power and the kind of explanatory insight they provide. I have also argued that Kahneman agrees with me that "RS explanations" constitute an important kind of non-causal scientific explanation. I expect that there is more to be learned about science from further research into RS explanations.[6]

Declarations

Conflict of interest There are no conflicts of interest to report.

References

John Beatty. (2006). Chance Variation: Darwin on Orchids. *Philosophy of Science* 73(5): 629–641.

Charles Darwin. (1872). *The Origin of Species*, 6th ed. London: John Murray.

Charles Darwin. (1883). *On the Variation of Animals and Plants Under Domestication*, volume 2. New York: Appleton.

Frederick Gravetter and Larry Wallnau. (2009). *Statistics for the Behavioral Sciences* (8th ed.). Belmont, CA: Wadsworth.

Ian Hacking. (1990). *The Taming of Chance*. Cambridge: Cambridge University Press.

Daniel Kahneman. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Marc Lange. (2013). Really Statistical Explanations and Genetic Drift. *PhilosophyofScience* 80: 169–188.

Marc Lange. (2017). *Because Without Cause: Non-Causal Explanations in Science and Mathematics*. New York: Oxford University Press.

Stefan Roski. (2021). In Defence of Explanatory Realism. *Synthese* 199: 14121–14141

Ernest Rutherford and Hans Geiger. (1910). The Probability Variations in the Distribution of α Particles. *Philosophical Magazine*, ser. 6, 20:698–707.

Elanor Taylor. (2018). Against Explanatory Realism. *Philosophical Studies* 175:197–219.

---

[6] My thanks to some anonymous *Synthese* referees who helped me to improve this paper.