Philosophy of Science (2022), **89**, 360–377
doi:10.1017/psa.2021.26

PSA PHILOSOPHY OF SCIENCE ASSOCIATION

PHILOSOPHY OF SCIENCE

ARTICLE

Resolving the Raven Paradox: Simple Random Sampling, Stratified Random Sampling, and Inference to Best Explanation

Barry Ward

Department of Philosophy, University of Arkansas, Fayetteville, AR, US
Email: bmward@uark.edu

(Received 17 April 2020; revised 03 September 2020; accepted 28 January 2021; first published online 12 January 2022)

Abstract

Simple random-sampling resolutions of the raven paradox relevantly diverge from scientific practice. We develop a stratified random-sampling model, yielding a better fit and apparently rehabilitating simple random sampling as a legitimate idealization. However, neither accommodates a second concern, the objection from potential bias. We develop a third model that crucially invokes causal considerations, yielding a novel resolution that handles both concerns. This approach resembles Inference to the Best Explanation (IBE) and relates the generalization's confirmation to confirmation of an associated law. We give it an objective Bayesian formalization and discuss the compatibility of Bayesianism and IBE.

1. The paradox

C. G. Hempel's (1945) raven paradox derives from a couple of prima facie plausible assumptions[1]:

> Nicod Criterion (NC): A claim of form "All Fs are Gs" is confirmed by any sentence of the form "i is F and i is G" where "i" is a name of some particular object.
>
> Equivalence Condition (EC): Whatever confirms (disconfirms) one of two equivalent sentences also confirms (disconfirms) the other.[2]

Taking the equivalent sentences (x)(Rx ⊃ Bx) and (x)(-Bx ⊃ -Rx) as, respectively, translating "all ravens are black" and "all nonblack things are nonravens,"[3] notoriously, the datum that an object is a nonblack nonraven confirms "all ravens are black."

[1] Hosiasson-Lindenbaum's (1940) discussion of the paradox predates Hempel's but attributes the paradox to him.

[2] By "equivalent" we mean classically logically equivalent. We shall not discuss resolutions that employ nonclassical logics.

[3] Stove (1966) rejects these translations. Such maneuvers are beyond our scope here.

© The Author(s), 2022. Published by Cambridge University Press on behalf of the Philosophy of Science Association.

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

The paradox has generated a variety of resolutions, but the dominant contemporary paradigm is a Bayesian one that models the generalization's confirmation in terms of simple random sampling. We offer two objections to this resolution. Our first, the objection from scientific practice, concerns striking disparities between simple random sampling and confirmational practice. This motivates the development of a stratified random-sampling model. It ultimately appears to vindicate simple random sampling as a legitimate idealization: the stratified approach provides a better fit with scientific practice but delivers essentially the same resolution. However, neither fares well against our second objection, the objection from potential bias.

This motivates a different approach, one that retains stratification, but on which we ultimately confirm the generalization by confirming a causal claim that entails it. Moreover, the appropriate methodology is not random sampling, but strongly resembles Mill's method of agreement (1868). Thus, it provides a very different resolution. In addition to handling both objections, it provides a better fit with scientific practice and hence an all-round better resolution than random-sampling approaches. It also compares favorably with Lipton's resolution (2004), which invokes causal considerations but identifies something much like Mill's method of difference as providing the key to the paradox.

Finally, we note some suggestive connections between the approach and Inference to the Best Explanation (IBE). We give the approach an objective Bayesian formalization and briefly discuss the compatibility of IBE and Bayesianism.

2. Simple random-sampling resolutions

Bayesians model confirmation of a hypothesis, h, by evidence, e, relative to background beliefs, K, using Bayes's theorem:

$$P(h|e \,\&\, K) = P(e|h \,\&\, K).\ P(h|K)/P(e|K)$$

On learning e, an agent updates her personal probability in h from her prior probability, P(h|K), to her posterior probability, P(h|e & K). Thus, e *incrementally confirms* h or simply *confirms* h, relative to K, if and only if $P(h|e \,\&\, K) > P(h|K)$. We may also speak of the *absolute* degree of confirmation of h, which is just h's probability. We shall say that h is *absolutely well confirmed*, or just *well confirmed*, for an agent if she gives h a high probability. What counts as a high probability is inevitably vague and none the worse for that. Much of our initial discussion concerns incremental confirmation, but well confirming shall become salient later.

Standardly, Bayesians deflate the paradox by arguing that Hempel's (1945) counterintuitive confirmers provide negligible incremental confirmation of the raven generalization. To obtain this result they assume our evidence is obtained by *simple* random sampling from the universe; i.e., it is assumed that each object in the population has an equal probability of being sampled. K is then specified to include reasonably plausible assumptions about the relative frequencies of nonblack things and ravens. Together, K and the simple random-sampling assumption mandate probabilities that resolve the paradox.

This approach originates with Hosiasson-Lindenbaum (1940) and has two chief variations. Comparative resolutions argue that a nonparadoxical instance, Ra & Ba, will confirm

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

significantly more than a paradoxical one, -Ra & -Ba. Noncomparative, or quantitative, resolutions argue that the confirmation afforded by -Ra & -Ba is positive, but small. Both claims presuppose some measure of confirmation. A reasonable and popular choice here is the difference between the posterior and prior probabilities of h, P(h|e & K) – P(h|K). However, the relevant results hold for a variety of such measures.[4]

Comparative resolutions commonly take K to justify a distribution, for which (a) P(-Ba|K) > P(Ra|K), (b) P(Ra|(x)(Rx ⊃ Bx) & K) = P(Ra|K), and (c) P(-Ba|(x)(Rx ⊃ Bx) & K) = P(-Ba|K). These accounts typically entail that nonblack nonravens are indeed confirming, although Fitelson and Hawthorne's (2010) recent account renders confirmation by black ravens greater than that afforded by nonblack nonravens without that seemingly gratuitous entailment. Quantitative resolutions depend on (c) and a strengthening of (a), P(-Ba|K) >> P(Ra|K).[5]

However, we're not concerned with the details of such proposals. We're concerned with two objections that apply generally.

3. The objection from scientific practice

The standard resolutions share two assumptions that starkly diverge from scientific practice: (i) evidence is gathered by simple random sampling and (ii) the only salient evidence concerns whether or not something is a raven and whether or not it is black. For instance, given the background belief that Arctic climes have caused white variants of many typically nonwhite species, it's quite plausible that *Arctic* ravens, if such there be, are also white. Thus, a sensible researcher will specifically seek them out to check that they are indeed black. More generally, she will be particularly interested in raven populations associated with distinctive conditions that might not implausibly be relevant to color and have little interest in mundane ravens of a kind that have already been frequently and consistently observed to be black. Thus, reasonable practice is incompatible with (i). It's also incompatible with (ii). For a reasonable researcher a black Arctic raven's confirmational significance should differ from that of a non-Arctic black raven of a kind with which she is already familiar—the magnitude of incremental confirmation should crucially depend upon such considerations. However, since the random sampler only conditionalizes on data regarding whether or not something is a raven or black, she does not even recognize being an *Arctic* raven as salient. Moreover, resolutions where one randomly samples from the set of ravens are equally culpable.[6] Let's call this the Objection from Scientific Practice.

What this objection plausibly shows is that our practice more closely resembles stratified random sampling. In stratified sampling we divide our population into mutually exclusive and jointly exhaustive relevantly homogeneous subpopulations and then randomly sample each as a means to evaluate our hypothesis regarding

[4] For details, see Fitelson and Hawthorne (2010). Fitelson (1999) discusses such measures more generally.

[5] There are variations, but these constraints, or close relatives, are commonly used. Fitelson and Hawthorne (2010) review both types of accounts. Vranas (2004) provides detailed discussion of quantitative ones.

[6] See, for instance, Suppes (1966), Gaifman (1979), and Horwich (1982).

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

the entire population.[7] There's good reason for a researcher to favor such a methodology. A stratified sampler who has already observed a substantial number of consistently black British ravens will get ever diminishing incremental confirmation of the generalization from further observations of (sigh!) *yet another* black British raven and will sensibly direct her research to other strata for which she currently has less data. By contrast, the simple sampler makes no such discriminations and will obliviously continue to sample from subpopulations that, were she pursuing the stratified approach, she could for all practical purposes ignore. So, the stratified sampler makes more efficient use of her primary epistemic resources, time and effort.

Notwithstanding the better fit with scientific practice, this doesn't really undermine the standard resolutions. Random-sampling resolutions are generally understood as idealizations of scientific practice, and the scientist's division of cases into kinds to be individually researched might be legitimately neglected in resolving the paradox. After all, if a researcher specifically randomly samples Arctic environments, the kinds of priors invoked by the simple random-sampling model should equally apply to that subpopulation, and hence the confirmation afforded by black Arctic ravens will be significantly greater than that afforded by nonblack nonravens drawn from the Arctic *for the same reasons*. The same will hold mutatis mutandis for random sampling of other strata. So, we have essentially the same comparative resolution of the paradox. The quantitative resolution should similarly apply. For a more precise treatment, see the appendix.[8]

4. The potential bias objection

Simple random sampling survives our first objection. There is, however, a further problem. For a sample to be random each member of the population must have an equal probability of being sampled. However, since we're interested in whether *all* ravens are black, long deceased and yet unborn ones are part of the relevant population and have zero probability of appearing in our sample. So, we can't *literally* randomly sample the relevant population.

If we've good reason to think the contemporary population is representative of, or suitably resembles, the total population past, present, and future, we can justify sampling the former as an unbiased sampling procedure for the latter. However, that's far from obvious. For all we know, past and future raven populations may be subject to selectional factors that cause nonblack ravens but are not represented in the contemporary population. Numerous causal mechanisms might yield such ravens. For instance, the plumage color in a bird population may change as a result of Batesian mimicry—the bird's coloration adapts to mimic the warning coloration of some local species that is, for instance, toxic to predators.[9] Indeed, more mundane cases are potentially problematic. As a matter of contingent fact, there might only be noncontemporary Arctic ravens or desert ravens, and so on. So, at least in the

[7] This commonly used statistical methodology is often employed in assessing voter preference, for instance. See Stuart (1962) for one exposition.

[8] In discussing consilience, Thagard (1978, 84) observes that a stratified approach should be favored for "all ravens are black" but does not develop such a model. To our knowledge, a stratified approach has not been pursued in the literature.

[9] For a case of Batesian mimicry in birds, see Londono, Garcia, and Martinez (2015).

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

initial stages of our research, we're not justified in holding sampling the contemporary population as an unbiased procedure for the total population. Moreover, it's not obvious we can ultimately justify this claim. Let's call this the Potential Bias Objection.

Could simple random sampling, in itself, address this? Might we justify holding our sample as unbiased, merely by sampling the contemporary population? Such sampling will indeed cover Arctic regions, for instance, along with the rest of the planet. However, since it only involves conditionalization on a sampled object's being a raven/nonraven and being black/nonblack, it does not register whether our sample includes Arctic ravens. So, it cannot tell us that the contemporary population is representative and hence that our sampling is indeed unbiased.

We could attempt to address this by simply taking note of the additional data, recognizing observed ravens as *Arctic* ravens, *desert* ravens, and so forth. Here's the rub. The data that would justify the judgement of representativeness, on its own, suffice to well confirm "*being a raven or something necessarily associated with that* causes blackness" by an application of Mill's method of agreement:

> If two or more instances of the phenomenon under investigation have one circumstance in common, the circumstance in which alone all the instances agree, is the cause (or effect) of the given phenomenon. (Mill 1868, 428)

Provided we're willing to countenance the possibility that being a raven, or something necessarily associated with that, causes blackness—and we are—then to the extent that we can gather data that ravens that are the product of all of the factors that might reasonably be causally relevant to raven color are all black, we have data that justify high confidence in the above causal claim. And since it entails "all ravens are black," the latter's probability must be at least as high. Thus, well confirming the causal claim by observation of a suitable variety of black ravens alone well confirms the generalization. So, meeting the rational precondition for random sampling by observing a suitable variety of ravens eliminates the need for such sampling and the need to attend to nonravens at all. We just go looking for Arctic ravens, desert ravens, and so forth.

Moreover, if we can't meet that precondition, random sampling is not merely redundant, it's nonviable. If, as a matter of contingent fact, there are no contemporary Arctic ravens to be observed, it's "game over" for the random sampler. Not for the causal confirmer, however. Unlike the random sampler, but like real scientists, she's not restricted to passive observation of a preexisting population. She can simply introduce raven populations to the Arctic and see how they fare. The experiment might be short and brutal, if ravens are indeed ill-adapted for Arctic survival. In that case, she confirms the causal claim by confirming that Arctic ravens are nomologically impossible: Arctic environments cannot cause nonblack ravens because they cannot sustain multigenerational raven populations. On the other hand, it might be long-term, if raven populations can indeed survive long enough for color selection to potentially manifest. Either way, she can well confirm the causal claim and hence the associated generalization.

So, simple random sampling is either redundant or nonviable. Indeed, nonviability seems the likelier of the two. A broad range of factors might be causally relevant to raven color—Arctic environments, desert environments, and a host of idiosyncratic

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

environmental pressures—and we certainly should not *expect* data restricted to the contemporary population to evade bias in general.

Stratified random sampling fares no better. If there happen to be contemporary exemplars of ravens for each stratum, then since the strata are, by hypothesis, relevantly homogeneous, random sampling of individual strata is redundant; observation of a modest number of black Arctic ravens should convince us that all Arctic ravens are black. And if any strata lack contemporary exemplars, it's nonviable. So, confirmation of the causal generalization and hence the entailed generalization by the method of agreement should be the preferred research methodology.

5. The causal resolution

We need to make the causal resolution explicit. In general, what are the salient confirming evidence statements for "being a raven or something necessarily associated with that causes blackness"? We shall need evidence that ravens that are the product of the potentially relevant causal factor are black, i.e., evidence of the form "Ra & Ba & Xa," where Xa states that a is a product of the potentially relevant factor. If ravens of kind X are to be found, then we may acquire it by passive observation. If not, we acquire it by experimentation. If we try and fail to produce multigenerational populations of Arctic ravens, say, the evidence statements will still concern *black ravens*, just dead ones.

This opens the door to a rather fast resolution of the paradox. We confirm the causal claim by *seeking* the salient data, by which we mean: we pursue that kind of data, adjust our probabilities by conditionalizing on it when we find it, and don't adjust our probabilities when we don't. A researcher who randomly samples from the raven population, as in the models noted in footnote 6, provides an example: she only conditionalizes on data furnished by ravens and ignores nonravens. The causal confirmer will seek data regarding Arctic ravens by looking in places where such ravens might be caused, i.e., the Arctic. Whether she finds naturally occurring Arctic ravens or is compelled to introduce them herself, and whether the introduced ravens survive or not, the salient confirming evidence concerns black ravens.[10] Since evidence of the form -Ra & -Ba is not salient to the causal claim, she does not conditionalize on it, and hence Hempel's (1945) paradoxical conclusion is false.[11]

The reason his argument is unsound is that the Nicod criterion is false: positive instances of "all nonblack things are nonravens" do not generally confirm it and its equivalents because we ignore them. Moreover, the considerations that dictate the data that are confirming are not syntactic. They ultimately depend upon the character of the causal claim whose confirmation confirms the generalization, the associated salient evidence, and our rational practice of seeking only salient evidence.

Can it be rational for an agent to *unapologetically* ignore the paradoxical evidence? The evidence that something is a nonblack nonraven has no obvious bearing on the

[10] If she finds nonblack ravens, such evidence is of course also salient, but disconfirming.

[11] Scheffler (1963, 284–85) observed that while, intuitively, -Ra & -Ba does not confirm "all ravens are black," it intuitively confirms the equivalent, "all nonblack things are nonravens." We might exploit this to raise trouble for some resolutions that do not suitably vindicate Scheffler's intuition. Not this one, however: -Ra & -Ba is manifestly not salient to the causal claim's confirmation.

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

causal hypothesis, and that seems reason enough. The scientist already knows how to do elementary causal inference, and she's not required to justify discounting prima facie irrelevant data she might happen across. If that were a general requirement, we'd all be irrational. Thus, the notion of rationality invoked above incorporates a type of means-end rationality, where the (immediate) end is well confirming the causal generalization if true and disconfirming it if false, and an *efficient* means to that end involves ignoring the potential import of observations regarding nonravens.

Certainly, every unexamined object is, in some sense, a potential disconfirmer. So, other things being equal, conditionalizing on the data that the rock you just tripped over is a (nonblack) nonraven would be marginally confirming, since those data entail that it's not a disconfirmer.[12] But even an agent who recognizes this recondite point—one who, presumably, does not see Hempel's conclusion as paradoxical—does not have good reason to conditionalize.[13] A policy of evaluating the import of such data would be a ludicrous expenditure of limited time, attention, and cognitive effort. Generally pursued, it would prevent us from gathering the data that could well confirm the generalization.

Our favored resolution is one on which the agent ignores data of the form "-Ra & -Ba," if she even tokens thoughts of such statements.[14] However, we don't have to dig our heels in here. Like proponents of standard random-sampling resolutions, we demand only that an adequate resolution explains our intuitions; it need not vindicate them. Thus, whether the paradoxical data we stumble across are universally ignored or occasionally conditionalized on is an incidental detail. What's distinctive about our resolution is that we confirm the generalization by confirming an associated causal claim by the method of agreement and that our data gathering proceeds by seeking the salient data, not by random sampling.

6. Formalizing the resolution

We now give our resolution a Bayesian formalization. We've already covered the paradoxical evidence statements. However, we must also formally characterize the confirmation afforded by data furnished by black ravens of various kinds.

[12] If an agent were to conditionalize, the confirmation afforded by e = -Ra & -Ba would be determined by the factor, $P(e|h)/(P(e) = P(e|h)/[(P(e|h)P(h) + P(e|-h)p(-h)] = 1/[p(h) + \{p(e|-h)/p(e|h)\}p(-h)]$, with e confirming h if the denominator <1. Since p(h) + p(-h) = 1, the denominator <1 if and only if $p(e|-h)/p(e|h) < 1$. Plausibly, P(-Ra & -Ba|-h) < P(-Ra & -Ba|h), since if h is true, there are no nonblack ravens, and so we should assign a probability of 1 that anything we stumble across is (-Ra & -Ba) v –(Ra & Ba) v (Ra & Ba), whereas if -h is true, we will presumably assign that a probability that is *marginally* less than 1. So, absent good reason to discriminate against stumbling across something that is -Ra & -Ba, some of that marginal "extra" probability should go to that kind of case. Thus, -Ra & -Ba would plausibly be weakly confirming if an agent were, implausibly, both competent and willing to make these rather involved estimates on the hoof.

[13] We should note that an agent who registers data but doesn't conditionalize on it is thereby vulnerable to a diachronic Dutch book argument. We give such vulnerability brief discussion in section 8.

[14] On Rinard's (2014) random-sampling resolution, nonravens also neither confirm nor disconfirm the generalization, but *not* because we ignore them. This striking result is ultimately derived from holding the kind *raven* more natural than the kind *black*. Our concerns with random sampling equally apply to Rinard's account. However, there's a loose but suggestive analogy between our respective appeals to causal considerations and to *comparative* naturalness of kinds.

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

We shall call these kinds *causally individuated kinds of Potential Disconfirmers* (hereafter, "kinds of PDs"). They meet two conditions. First, they are kinds of cases individuated by factors that an agent holds might reasonably be causally relevant to the production of disconfirming cases. Thus, corresponding to each kind of PD there is a background *belief* that an associated factor could, reasonably, be a cause of disconfirming cases. Second, to count as a kind of PD in the intended sense—i.e., one that can facilitate well confirming by merely observing a modest number of exclusively black ravens of that kind—she must have *a commitment* that there are no additional factors that might cause nonblack ravens of that kind: *Arctic raven* doesn't count as a kind of PD unless she has a commitment that there are no selectional factors other than being an Arctic raven that might cause nonblack Arctic ravens. Identifying that commitment as a background belief and hence assigning it probability 1 is too demanding. More reasonably, we demand only *a sufficiently high degree of confidence* that there are no such factors. If the agent lacks an appropriate confidence, then either her set of kinds should be more fine-grained—if she can specify the further factors—or she's too ignorant by her own lights to rationally well confirm "being an Arctic raven or something necessarily associated with that causes black plumage."[15] However, granted the required confidence, observing a modest number of black Arctic ravens should well confirm the blackness of Arctic ravens.

Formally, let background beliefs, K, determine that K_i is a kind of case for which we have a sufficiently high degree of confidence that there are no additional factors causally relevant to property B, and let C_i be the corresponding causal generalization, "being an entity of kind K_i or something necessarily associated with that causes the entity to be B." We impose a constraint on our agent's priors as follows. Define $P^* = P(.|K)$. Then, if $e_1, e_2, \ldots, e_n$ consists of confirming data of the corresponding kind; i.e., each is a statement of the form, "$K_i a$ & Ba," then $P^*(C_i| e_1 \& e_2 \& \ldots \& e_n)$ must be high/well confirmed, for even modest values of n.[16]

We can view this constraint as embodying a particularly simple application of the method of agreement. The method demands a prior list of potential causes we are willing to countenance: since there is an arbitrarily large set of circumstances in which a finite list of objects agree, there can be no inference to "the circumstance in which alone all the instances agree" without such a list. Our agent's high degree of confidence that there are no additional factors causally relevant to property B in effect reduces her list of candidates to one entry—being a case of kind K_i or something necessarily associated with that. Hence, observation of a sequence of cases of kind K_i that are all Bs demands she well confirms that "being an entity of kind K_i or something necessarily associated with that causes the entity to be B," and the constraint on priors enforces that.

So, let K_i be *Arctic ravens* and C_i the corresponding causal generalization, "being an Arctic raven or something necessarily associated with that causes

[15] Notoriously, Europeans were surprised to encounter black Australian swans. Their high antecedent confidence that all swans are white suggests unjustified confidence in their knowledge of factors that might be causally relevant to swan color. Although, even the most conscientious researchers may be surprised.

[16] But not equal to 1. We don't wish to preclude future disconfirmation of the generalization by conditionalization.

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

black plumage." Conditionalizing on data of form "$K_i a$ & Ba" furnished by a modest number of Arctic ravens will render $P^*(C_i)$ well confirmed. Further, let R_i be the corresponding generalization, "all Arctic ravens are black." Since C_i entails R_i, $P^*(R_i | e_1 \,\&\, e_2 \,\&\, \ldots \,\&\, e_n) \geq P^*(C_i | e_1 \,\&\, e_2 \,\&\, \ldots \,\&\, e_n)$. So, such priors dictate that R_i is well confirmed by the corresponding observations of black Arctic ravens.

Of course, merely well confirming the individual generalizations associated with each kind need not well confirm, or even substantially confirm, "all ravens are black," which is equivalent to their conjunction. It's not merely coherent, but often very reasonable to have a high probability for each of a set of conjuncts and a low probability for their conjunction. On the other hand, if she *sufficiently well confirms* that each kind could not cause a nonblack raven, the rational agent effectively eliminates all of the explicit reasons she has for thinking that each particular environment might cause nonblack ravens.[17] Given that she has eliminated all of these potential, more parochial explanations of the color of raven plumage, she should certainly countenance the hypothesis, "being a raven or something necessarily associated with that causes blackness." And at that point, with all due respect to inductive humility, if she has done due diligence in identifying factors that might have been causally relevant to disconfirmation, she has reasonably met the conditions for the overarching application of Mill's method of agreement (1868), one that takes the evidence provided by all the observations of black ravens from her set of kinds as well confirming "being a raven or something necessarily associated with that causes black plumage." Thus, it will be inductively rational for her to hold "being a raven or something necessarily associated with that causes black plumage" and hence the entailed generalization well confirmed.

We should acknowledge that the great variety of factors that might play a role in natural selection could render due diligence very demanding. However, that merely reflects the reality of scientific research: well confirming a general hypothesis typically demands a wealth of causal knowledge. Individuals may propose such hypotheses, but their confirmation often involves the accumulation of such knowledge by large numbers of scientists over decades or even centuries. On the other hand, it need not be a practical impossibility. In any case, let's assume she can do due diligence and rationally well confirm the general causal claim. How should we model her epistemic evolution?

It might seem tempting to employ some further constraint on priors. However, that would be misguided. Our agent individuated Arctic ravens as a kind of PD, precisely because she *believed* that some factor specifically associated with that kind might cause nonblack plumage. In sufficiently well confirming that Arctic ravens are black, she reasonably comes to reject that belief, and the same goes, mutatis mutandis for the beliefs associated with other kinds of PDs.[18] Thus, her transformation involves a

[17] We now have two distinct notions of "sufficiently well confirmed" in play. The first is sufficiently well confirmed *that there are no further factors relevant to a given kind of case that we can reasonably impose our Millian constraint on priors*. The second is the one just introduced.

[18] We're not claiming there's some threshold of degree of belief that in all epistemic contexts suffices for belief. In this case, however, while allowing that such a "threshold" may be a vague interval and may permissibly vary amongst individuals, we do hold that a reasonable agent will have one. Given suitable evidence, retaining the belief that being an Arctic raven provides a reason to suspect disconfirmation amounts to an irrational degree of inductive humility. From a practical point of view, she moves below her threshold when Arctic ravens cease to be a reasonable focus of research for this hypothesis.

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

revision of background beliefs, in which she eliminates these kinds of PDs. At that point, *raven* is the narrowest kind of potentially disconfirming case that she individuates.

Should we count *raven* as a kind of PD, in our technical sense of that term? No, it's not that she's identified some specific factor associated with ravens that might be causally relevant to the production of nonblack ravens. Quite the opposite: having done due diligence, she's highly confident there are none. At this point, should she *believe* that "being a raven or something necessarily associated with that causes black plumage" and the entailed generalization? Inductive humility likely counsels against that. However, the large number of black ravens that she has observed—the ones that furnished the data she used to eliminate her initial set of kinds of PDs—should well confirm "being a raven or something necessarily associated with that causes black plumage" and hence the generalization.

Let's flesh that out formally. Her old distribution had the background belief, "being an Arctic raven provides a reason to suspect disconfirmation," and so assigned that proposition probability 1. Her new distribution drops that belief and indeed should assign the proposition a low probability. Hence, her new distribution cannot be obtained by conditionalization on the old.[19] Second, in addition to dropping the beliefs that individuated her initial set of kinds of PDs, her new set of background beliefs, K, includes the belief that she has done due diligence for the kind, *raven*; i.e., it assigns a high probability to there being no factors that might cause nonblack ravens. Let's call this new distribution, $P^{**}(.) = P(.|K)$. Since it assigns a high probability to there being no factors that might cause nonblack ravens, it should also conform to our Millian constraint on priors; i.e., given that C is "being a raven or something necessarily associated with that causes black plumage," the constraint dictates that $P^{**}(C|\ e_1\ \&\ e_2\ \&\ \ldots\ \&\ e_n)$ has a high value for even modest values of n, where the e_j statements are now simply of the form "Ra & Ba." Since her numerous prior observations of black Arctic ravens, black desert ravens, and so forth provided her with a large number of such data reports, P** should assign each of those probability 1. So, there is a more than modest conjunction of such statements, for which $P^{**}(e_1\ \&\ e_2\ \&\ \ldots\ \&\ e_n) = 1$. Hence, P**(C) must be high/well confirmed and hence also P**(R), where R = "all ravens are black." Her new distribution assigns both claims high probability.

Let's be the first to acknowledge that we haven't specified a rule that uniquely determines the new distribution. There may not be such a rule. What we're happy to defend is that her rational evolution is constrained by the above considerations.

7. Is the causal resolution idiosyncratic to the raven generalization?

We argued that the sampled raven population may not be representative because it does not include past and future ravens. This might suggest that potential bias is not a concern for more typical scientific generalizations. Not so, however.

Consider another Hempelian example, "all sodium salts burn yellow." The absence of historically contingent selection effects on the chemical character of sodium salts makes it reasonable to suppose that contemporary sodium salts are indeed

[19] If $p(A) = 1$, then $p(A|B) = 1$ for arbitrary B, provided that $P(B) > 0$.

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

representative of all. However, unlike ravens, they're not restricted to the planet Earth, or indeed our local group of galaxies. So, again, we cannot literally randomly sample the population of contemporary sodium salts or any population that includes them. As before, to justify the judgment that our sample is representative we would have to ensure that it includes cases of each chemical kind of sodium salt whose specific character might give us reason to suspect disconfirmation and, indeed, any variations in crystal structure or other allotropic factors that might provide such reason. So, as before, meeting the rational precondition for random-sampling demands we observe a suitable variety of sodium salts, which will suffice to well confirm the causal generalization and hence the entailed generalization, rendering random sampling redundant and giving us no reason to pay attention to nonsodium salts. And if, as will often be the case, finding samples of a given kind of sodium salt is inefficient or a practical impossibility, we do what scientists do on a daily basis, proceed by experimentation, i.e., attempt to synthesize the relevant chemical. Here again, the causal approach is preferable.

Its broader applicability might prompt the suspicion that we've shown too much: that we can only randomly sample in situations where we can more readily confirm an associated causal claim. However, that doesn't follow. In cases where the total population or suitable strata of that population can literally be randomly sampled, as in an election poll, quality control at a widget factory, or a survey of household income, there need be no concerns about representativeness and no attendant confirmation of an associated causal claim. Moreover, there are countless cases where literal random sampling is impossible, but potential bias is not a concern because background knowledge allows us to reasonably expect the total population to be relevantly similar to the sampled one. If we think the distribution of factors causally relevant to the incidence of pancreatic cancer in the contemporary population will remain constant for future populations, we will happily use random sampling of the contemporary population as a guide to incidence in future populations and total populations across time.

We should acknowledge the daunting prospect of well confirming that an almost limitless variety of kinds of sodium salts all burn yellow. If each allotrope of each chemical kind of sodium salt is a kind of PD, such a research project seems unmanageable. However, the consolidation of kinds manifested in our formal treatment of the raven applies here at multiple levels. Observation of a sufficient variety of yellow-burning allotropes of sodium halides may cause us to consolidate allotropes of sodium chloride, sodium fluoride, and so forth into one kind of PD, *sodium halide*, eliminating the need to observe each allotropic and chemical kind in that family; uniform yellow burning among a variety of other kinds of PDs that we have some initial reason to suspect may behave similarly can prompt other consolidations, until ultimately we are left with one kind of case, *sodium salt*. An initially intractable research program evolves, through unification of kinds, into a manageable one.[20]

Whether we find or synthesize a given kind of sodium salt, the confirming evidence statement will be of the form Sa & Ya & Xa, where X specifies the particular factor that might be causally relevant to color of burning. If we fail to find a sodium salt by

[20] This kind of consolidation might be related to a notion of dynamical consilience, although not one subsumed by the model in Thagard (1978).

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

seeking, we will, as before, reasonably ignore the data. And if we ultimately determine the nomological impossibility of some putative kind of sodium salt, the evidence statements that confirm impossibility and hence the causal generalization will not be of the form "-Sa & -Ya." Chemists may repeat an experiment many times and try a variety of different ways of synthesizing a compound or crystal structure before succeeding. Correspondingly, judging that a kind is nomologically impossible will often depend upon multiple strands of evidence, each with its idiosyncratic specification of the relevant conditions. If we fail to synthesize a theoretically postulated novel crystalline form of a sodium salt, the relevant evidence statement might be "The sample, a, obtained by crystallization with gradual cooling of the supersaturated solution was a pure sodium salt, but did not have the hypothesized crystal structure." Or if the product is not a sodium salt, the color with which it burns will be irrelevant. Hempel's (1945) paradoxical conclusion is false here too because we are confirming the generalization by confirming causal claims.

8. Causal approaches, law confirmation, and inference to the best explanation

We should contrast our resolution with Lipton's (2004, chapter 6), which, like ours, depends upon causal considerations. Unlike us, Lipton does not focus on confirmation of the generalization. Rather, he treats the raven hypothesis as a causal claim:

> . . . there is something in ravens, a gene perhaps, that makes them black. Moreover, the hypothesis implies that this cause, whatever it is, is part of the essence of ravens (or at least nomologically linked to their essence). (Lipton 2004, 97)

He then explicates its confirmation in terms of contrastive inference (chapter 5), a close relative of Mill's (1868) method of difference. Some contrapositive instances of a causal claim, those that have a similar causal history to its positive instances, will provide evidence for the claim. Thus, for Lipton, a non-yellow-burning flame, in which no sodium salt is present (relevant contrapositive), plus the yellow flame obtained upon introducing a sodium salt (contrasting positive) together support "all sodium salts burn with a yellow flame," construed as a causal generalization. However, typical contrapositives—a white shoe, say—are not relevantly similar to any positive instance and provide no support. Hence, Hempel's (1945) paradoxical conclusion is false.

Since our methodology closely resembles Mill's method of agreement, whereas Lipton's is a method of difference, they fundamentally differ. Of course, our account primarily concerns confirming the generalization. Confirming the causal generalization is a means to that end, but that doesn't explain the difference. There's a mistake in Lipton's analysis. Given that we antecedently know, as we do, that each sodium salt's constitution causes the color with which it burns, even contrapositive instances that are suitably similar to the positive ones are not directly relevant to whether *some aspect of its essence/something necessarily associated with being a sodium salt* causes yellow burning. They merely provide evidence against the claim that something necessarily associated with *some more inclusive kind than sodium salts* causes yellow burning. The same point applies mutatis mutandis to Lipton's construal of the raven hypothesis.

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

The method of difference is just not relevant to the confirmation of Lipton's causal claims, given reasonable background beliefs.[21] Similarly, when we seek to confirm the generalization, whether or not some broader generalization subsumes it, is beside the point.

While our primary concern is confirmation of the raven and sodium salt generalizations, we do so by well confirming a closely related law/broad causal generalization. Why is that? Both examples are generalizations, for which (i) we cannot confidently exclude any nomologically possible kind of PD from actuality, (ii) we confirm using our causal methodology, and (iii) we're willing to countenance the possibility of an associated unifying causal factor, e.g., *being a raven or something necessarily associated with that*. In both cases, given (i), well confirming the generalization demands we well confirm its nomological necessity, and (ii) and (iii) facilitate well confirming that necessity by well confirming an entailing law.

Now, we're not going to casually ascribe laws a sweepingly general role in confirmation. However, these two examples are not mere special cases. Given the spatiotemporal vastness and causal variety of our universe, we'll realize condition (i) for many generalizations that are *sufficiently informative* to serve the ambitious predictive goal of science.[22] Hence, to well confirm them we will need to well confirm their necessity. So, ones we think would likely be accidental if true, we will judge as poor candidates. By contrast, to the extent that we suspect an informative generalization holds as a matter of law, we suspect its necessity. Hence, notwithstanding (i), we have a reason to suspect its actual truth—by our lights, it has a fighting chance of being true. Moreover, it will satisfy (iii), allowing us to invoke the method of agreement to well confirm it, given data gathered in accord with (ii) for each kind of PD.[23] So, targeting suspected laws is one reasonable strategy for confirming informative generalizations for two reasons: it gives us reason to suspect the generalization is actually true and we have a methodology that can well confirm such generalizations. These are not minor virtues for an agent in our impoverished epistemic situation with major predictive ambitions. So, we can see why law confirmation could play an important role in our understanding of scientific confirmation more generally.

There's also a suggestive connection with Inference to the Best Explanation (IBE). We well confirm "all ravens are black" by well confirming the explanatory generalization, "being a raven or something necessarily associated with that causes black plumage," and decisively disconfirming the causal relevance of other factors and associated, more parochial, explanations of ravens' colors: in confirming that Arctic ravens are black, we confirm that being an *Arctic* raven is irrelevant to color and

[21] Notwithstanding his official resolution, Lipton (2004, 99) gives his method of difference no role in confirming the (causal) raven hypothesis itself: "Contrastive inference avoids the raven paradox, but it does not account for the way we support the raven hypothesis itself." So, even he recognizes his method of difference as, at least, inessential to the confirmation of such claims.

[22] *Not all*, of course. Reichenbach's generalization, "all solid spheres of gold have a diameter of less than 1 mile," is highly informative, but we judge it likely true, in part because we judge that nomologically possible kinds of cases where, for instance, colossal gold spheres form by gravitational accretion from the interstellar medium are likely nonactual.

[23] If the suspected law involves multiple causally relevant factors, it will not satisfy (iii), but will satisfy a similar condition and might be confirmed by something more along the lines of the method of concomitant variation. For obvious reasons, we're disregarding such subtleties here.

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

disconfirm associated explanatory hypotheses such as "being an *Arctic* raven or something necessarily associated with that causes it to be white" and similarly for other parochial explanations. Thus, we well confirm one explanatory hypothesis by eliminating multiple competitors from serious consideration.

Notoriously, IBE and Bayesianism are commonly held to make poor bedfellows. Given the prima facie reasonableness of the causal research strategy and its apparent fit with quotidian scientific methodology, we might be tempted to say, "so much the worse for Bayesianism." However, we have pursued a Bayesian account. Let's assess the marriage.

As per section 6, beliefs and confidences about possible causal dependencies are taken to impose rationality constraints on our priors. These constraints, which effectively enforce particular applications of Mill's method of agreement, are not suggestions. If you don't think that observations of a modest number of black Arctic ravens should well confirm that all Arctic ravens are black, *even though you are highly confident that no Arctic ravens differ in respects that are causally relevant to color*, then you're as close to inductive skepticism as be damned. You're certainly not a competent scientist. So, we reject the reconciliation of subjective Bayesianism and IBE advocated by Lipton and others, on which IBE is a mere heuristic that may be rationally rejected by an agent with conflicting priors.[24] Our view combines IBE and a species of objective Bayesianism. *Objective* here need not require that, given background beliefs, K, rationality demands conformity to some unique probability distribution P(.|K). However, distributions that violate the above synchronic constraints—and, as we can readily envision, others required to enforce other inductively rational species of causal/explanatory inferences—are not permissible for an inductively rational agent.

So, what are the problems? van Fraassen (1989, 166–9) has argued that an agent who implements IBE by following a rule that gives a probability boost to a hypothesis that best explains the data in addition to that provided by conditionalizing on the data is irrational by virtue of being vulnerable to a diachronic Dutch book. One component of our Bayesian story is a constraint on priors, not a distinct updating rule. So, we have no cause for concern there. The other component, where we introduce a new distribution that demotes some background beliefs, might expose us to a diachronic Dutch book. But, what of it? Once we acknowledge—as surely we must—the need to sometimes discard beliefs, we must reject conditionalization as a general policy for updating. And as Douven (2013, 430), among others, observes, even if we accept Dutch book arguments as specifying synchronic constraints on epistemic rationality, there's no inconsistency in holding conflicting views on the fairness of bets at different times. Indeed, such diachronic conflicts result from Bayesian conditionalization itself.

Roche and Sober's (2013; hereafter, R & S) problem invokes the subjunctive, E, "If H were true and O were true, then H would explain O." They argue that, in general, $P(H|O \,\&\, E) = P(H|O)$, i.e., O screens-off E from H, and conclude that H's explanatoriness is evidentially irrelevant. From there, they argue against the reconciliation of IBE and Bayesianism. Now, our constraint on priors is not in the requisite form for their argument. It demands that $P^*(C_i|\, e_1 \,\&\, e_2 \,\&\, \ldots \,\&\, e_n)$ is high where each e_j is a statement of

[24] Indeed, as Jonathan Weisberg (2009) argues, IBE can't survive a marriage with subjective Bayesianism, since freedom to pick priors is a license to ignore explanatory considerations.

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

the form "$K_i a_j$ & $B a_j$." However, what C_i explains is that, *given that* a_j *is a raven,* a_j is black, for each e_j. Or we could say that it explains the agreement between the cases, i.e., that all the observed ravens agree in being black. We're happy to retain our constraint in the form that directly matches the method of agreement, rather than pressing it into a form directly specifying the probability for C_i conditional on the evidence it explains, but of course that doesn't constitute a defense.

There have been several critical responses to R & S's argument. Let's just register our own, quite general, concern with screening-off as a criterion of evidential relevance. R & S motivate it using a prima facie reasonable Bayesian criterion for evidential relevance:

O is evidentially relevant to H if and only if $P(H|O) > P(H)$.

However, if P(.) is our current personal probability, and we have antecedently learned O—O is a background belief already codified in P(.)—then $P(H|O) = P(H)$, even if learning O rendered P(H) higher than *it would otherwise have been*. So, O can be evidentially relevant to H, even if the inequality does not hold. Certainly, once O has been codified in P(.) it has no *additional* relevance, but that's beside the point.

In R & S's case where we are considering the evidential relevance of one proposition to another conditional on taking a further proposition's relevance into account, the problem is particularly salient. For instance, suppose O, "the barometer reading is falling," is my evidence for F, "the air pressure is falling," which in turn is my evidence for H, "there will be a storm." Further, suppose that O is my sole evidence for F; i.e., I will believe F only if I believe O. In that case, $P(O|F) = 1$, and so, $P(H|F \ \& \ O) = P(H|F)$; i.e., F screens-off O from H. However, O is manifestly evidentially relevant to F and hence to H. Thus, screening-off is a poor criterion of evidential relevance.

So, even if C_i were screened off from the associated E by the data it explains, that would not show the evidential irrelevance of explanatoriness. In the barometer case, a counterfactual criterion provides a better rule of thumb for evidential relevance: if I hadn't learned O, then I wouldn't have learned, or at least raised my probability for, F or H. Similarly, if the causal relations specified by C_i *hadn't* been the best explanation of aspects of the data specified by e_1 & e_2 & . . . & e_n, inductive rationality, as embodied by Mill's method, *wouldn't have* demanded that it be strongly confirmed by that data. Whether or not some associated screening-off claim is true, C_i's explanatoriness is manifestly evidentially relevant.

9. Conclusion

We've provided good reason to reject both simple and stratified random-sampling resolutions of the paradox and developed our own causally driven resolution. There are also intimations of a broader confirmation theory that can cover generalizations, for which—as is often true in the sciences—random sampling is susceptible to potential bias. On this confirmation theory, numerous such generalizations may be confirmed by confirming an entailing explanatory law or causal generalization by what looks like a species of IBE. So, our resolution should not only be of interest to students of the paradox, but to confirmation theorists more generally and epistemologists with an interest in IBE.

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

Acknowledgements. I would like to thank Michael Goldsby and an audience at the 2017 Pacific Division meeting of the APA for helpful comments on an ancestor of this paper and the members of a University of Arkansas reading group and three anonymous referees for their helpful comments on a previous draft.

References

Douven, Igor. 2013. "Inference to the Best Explanation, Dutch Books, and Inaccuracy Minimisation." *Philosophical Quarterly* 63:428–44.
Fitelson, Branden. 1999. "The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity." *Philosophy of Science* 66 (Proceedings):S362–78.
Fitelson, Branden, and James Hawthorne. 2010. "How Bayesian Confirmation Theory Handles the Paradox of the Ravens." In *Probability in Science. Boston Studies in the Philosophy and History of Science*, vol. 284, ed. Ellery Eells and James Fetzer, 247–77. Dordrecht, Heidelberg, London, New York: Springer.
Gaifman, Haim. 1979. "Subjective Probability, Natural Predicates and Hempel's Ravens." *Erkenntnis* 14:105–47.
Hempel, Carl G. 1945. "Studies in the Logic of Confirmation (I)." *Mind* 54:1–26.
Horwich, Paul. 1982. *Probability and Evidence*. Cambridge: Cambridge University Press.
Hosiasson-Lindenbaum, Janina. 1940. "On Confirmation." *Journal of Symbolic Logic* 5:133–48.
Lipton, Peter. 2004. *Inference to the Best Explanation* (2nd Edition). New York: Routledge.
Londono, Gustavo A., Duvan A. Garcia, and Manuel A.S. Martinez. 2015. "Morphological and Behavioral Evidence of Batesian Mimicry in Nestlings of a Lowland Amazonian Bird." *The American Naturalist* 185:135–41.
Mill, John S. 1868. *A System of Logic, Ratiocinative and Inductive Vol. I* (7th Edition). London: Longmans, Green, and Co.
Roche, William, and Elliott Sober. 2013. "Explanatoriness Is Evidentially Irrelevant, or Inference to the Best Explanation Meets Bayesian Confirmation Theory." *Analysis* 73:659–68.
Rinard, Susanna. 2014. "A New Bayesian Solution to the Paradox of the Ravens." *Philosophy of Science* 81:81–100.
Scheffler, Israel. 1963. *The Anatomy of Inquiry: Philosophical Studies in the Theory of Science*. New York: Knopf.
Stove, David. 1966. "Hempel's Paradox." *Dialogue: Canadian Philosophical Review* 4:444–55.
Stuart, Alan. 1962. *Basic Ideas of Scientific Sampling*. London: Griffin.
Suppes, Patrick. 1966. "A Bayesian Approach to the Paradoxes of Confirmation." In *Aspects of Inductive Logic* ed. Jaakko Hintikka and Patrick Suppes, 198–207. Amsterdam: North Holland.
Thagard, Paul. 1978. "The Best Explanation: Criteria for Theory Choice." *The Journal of Philosophy* 75:76–92.
van Fraassen, Bas C. 1989. *Laws and Symmetry*. New York: Oxford University Press.
Vranas, Peter B.M. 2004. "Hempel's Raven Paradox: A Lacuna in the Standard Bayesian Solution." *The British Journal for the Philosophy of Science* 55:545–60.
Weisberg, Jonathan. 2009. "Locating IBE in the Bayesian Framework." *Synthese* 167:125–43.

Appendix: Stratified random sampling for (x)(Rx ⊃ Bx)

We divide the population into homogeneous strata: Arctic environments, desert environments, etc. We then confirm each corresponding generalization, "all Arctic ravens are black," "all desert ravens are black," and so forth, by randomly sampling each stratum. The data considered have either the form "Ra & Ba & S_ia" or "-Ra & -Ba & S_ia" where S_i = "is from stratum i." We shall show that both the comparative and quantitative resolutions follow, given two weak and reasonable additional assumptions.

Comparative resolution:
For each stratum, we make assumptions corresponding to the standard assumptions for simple random sampling (as per section 2):

(a′) $P(\text{-Ba} \,\&\, S_i a|K) > P(\text{Ra} \,\&\, S_i a|K)$
(b′) $P(\text{Ra} \,\&\, S_i a|(x)((\text{Rx} \,\&\, S_i x) \supset \text{Bx}) \,\&\, K) = P(\text{Ra} \,\&\, S_i a|K)$
(c′) $P(\text{-Ba} \,\&\, S_i a|(x)((\text{Rx} \,\&\, S_i x) \supset \text{Bx}) \,\&\, K) = P(\text{-Ba} \,\&\, S_i a|K)$

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

These guarantee that, using the difference measure of confirmation, P(h|e & K) – P(h|K), black ravens from each stratum confirm the corresponding generalization more than nonblack nonravens from that stratum. Writing P* = P(.|K), we have:

$P^*((x)((Rx \,\&\, S_ix) \supset Bx)|Ra \,\&\, Ba \,\&\, S_ia) - P^*((x)((Rx \,\&\, S_ix) \supset Bx)) > P^*((x)((Rx \,\&\, S_ix) \supset Bx)|\text{-}Ra \,\&\, \text{-}Ba \,\&\, S_ia) - P^*((x)((Rx \,\&\, S_ix) \supset Bx))$

Hence:

(i) $P^*((x)((Rx \,\&\, S_ix) \supset Bx)|Ra \,\&\, Ba \,\&\, S_ia) > P^*((x)((Rx \,\&\, S_ix) \supset Bx)|\text{-}Ra \,\&\, \text{-}Ba \,\&\, S_ia)$

We shall use (i) in deriving the comparative claim for the raven generalization itself, to which we now turn.

Given exhaustiveness, "all ravens are black" is equivalent to the conjunction of the generalizations corresponding to the strata:

$P^*((x)(Rx \supset Bx)) = P^*((x)((Rx \,\&\, S_1x) \supset Bx) \,\&\, \ldots \,\&\, (x)((Rx \,\&\, S_nx) \supset Bx))$, for the appropriate n.

For each i, we can write:

$P^*((x)(Rx \supset Bx)) = P^*((x)((Rx \,\&\, S_1x) \supset Bx) \,\&\, \ldots \&(x)((Rx \,\&\, S_{i-1}x) \supset Bx) \,\&(x)((Rx \,\&\, S_{i+1}x) \supset Bx) \&\, \ldots \,\&\, (x)((Rx \,\&\, S_nx) \supset Bx)|(x)((Rx \,\&\, S_ix) \supset Bx).\ P^*((x)((Rx \,\&\, S_ix) \supset Bx)))$.

Since, in general, P(A & B|C) = P(A| B & C). P(B|C),[25] we obtain:

$P^*((x)(Rx \supset Bx)|Ra \,\&\, Ba \,\&\, S_ia) = P^*((x)((Rx \,\&\, S_1x) \supset Bx) \,\&\, \ldots \,\&(x)((Rx \,\&\, S_{i-1}x) \supset Bx) \,\&(x)((Rx \,\&\, S_{i+1}x) \supset Bx) \&\, \ldots \,\&\, (x)((Rx \,\&\, S_nx) \supset Bx)|(x)((Rx \,\&\, S_ix) \supset Bx) \,\&\, Ra \,\&\, Ba \,\&\, S_ia).\ P^*((x)((Rx \,\&\, S_ix) \supset Bx)|Ra \,\&\, Ba \,\&\, S_ia)$

and

$P^*((x)(Rx \supset Bx)|\text{-}Ra \,\&\, \text{-}Ba \,\&\, S_ia) = P^*((x)((Rx \,\&\, S_1x) \supset Bx) \,\&\, \ldots \,\&(x)((Rx \,\&\, S_{i-1}x) \supset Bx) \,\&(x)((Rx \,\&\, S_{i+1}x) \supset Bx) \&\, \ldots \,\&\, (x)((Rx \,\&\, S_nx) \supset Bx)|(x)((Rx \,\&\, S_ix) \supset Bx) \,\&\, \text{-}Ra \,\&\, \text{-}Ba \,\&\, S_ia).\ P^*((x)((Rx \,\&\, S_ix) \supset Bx)|\text{-}Ra \,\&\, \text{-}Ba \,\&\, S_ia)$

since we have shown:

(i) $P^*((x)((Rx \,\&\, S_ix) \supset Bx)|Ra \,\&\, Ba \,\&\, S_ia) > P^*((x)((Rx \,\&\, S_ix) \supset Bx)|\text{-}Ra \,\&\, \text{-}Ba \,\&\, S_ia)$, for each i.

Using the difference measure, we can infer that black ravens drawn from any stratum confirm "all ravens are black" more than nonblack nonravens drawn from that stratum, granted the following two weak and reasonable assumptions:

(A1) Finding a black raven in stratum i does not decrease our probability for all ravens from other strata are black conditional on the claim that all ravens from stratum i are black.

(A2) Finding a nonblack nonraven in stratum i does not increase our probability for all ravens from other strata are black conditional on the claim that all ravens from stratum i are black.

[25] As we can see by expanding the conditional probabilities on each side of the equation. The right-hand side is [P(A & B & C)/P(B & C)].[P(B & C) /P(C)], which is just P(A & B & C)) /P(C), i.e., P(A & B|C).

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press

Respectively:

(A1) $P^*((x)((Rx \,\&\, S_1x) \supset Bx) \,\&\, \ldots \,\&(x)((Rx \,\&\, S_{i-1}x) \supset Bx) \,\&(x)((Rx \,\&\, S_{i+1}x) \supset Bx) \,\&\, \ldots \,\& (x)((Rx \,\&\, S_nx) \supset Bx) \mid (x)((Rx \,\&\, S_ix) \supset Bx) \,\&\, Ra \,\&\, Ba \,\&\, S_ia) \geq P^*((x)((Rx \,\&\, S_1x) \supset Bx) \,\&\, \ldots \,\&(x)((Rx \,\&\, S_{i-1}x) \supset Bx) \,\&(x)((Rx \,\&\, S_{i+1}x) \supset Bx) \,\&\, \ldots \,\& (x)((Rx \,\&\, S_nx) \supset Bx) \mid (x)((Rx \,\&\, S_ix) \supset Bx))$

and

(A2) $P^*((x)((Rx \,\&\, S_1x) \supset Bx) \,\&\, \ldots \,\&(x)((Rx \,\&\, S_{i-1}x) \supset Bx) \,\&(x)((Rx \,\&\, S_{i+1}x) \supset Bx) \,\&\, \ldots \,\& (x)((Rx \,\&\, S_nx) \supset Bx) \mid (x)((Rx \,\&\, S_ix) \supset Bx) \,\&\, \text{-}Ra \,\&\, \text{-}Ba \,\&\, S_ia) \leq P^*((x)((Rx \,\&\, S_1x) \supset Bx) \,\&\, \ldots \,\&(x)((Rx \,\&\, S_{i-1}x) \supset Bx) \,\&(x)((Rx \,\&\, S_{i+1}x) \supset Bx) \,\&\, \ldots \,\& (x)((Rx \,\&\, S_nx) \supset Bx) \mid (x)((Rx \,\&\, S_ix) \supset Bx))$.

Thus, we obtain the same comparative result regarding incremental confirmation of "all ravens are black" as we did for each of the stratified generalizations, essentially reproducing the resolution afforded by standard simple random-sampling resolutions.

Quantitative resolution:

If we are pursuing the quantitative approach, we use (c′) and (a′′), $P(\text{-}Ba \,\&\, S_ia|K) >> P(Ra \,\&\, S_ia|K)]$. For ease of reading, let $P^{**} = P(.|K)$ for these background assumptions, K.

This will reproduce the quantitative result for the generalization corresponding to each stratum:

$P^{**}((x)((Rx \,\&\, S_ix) \supset Bx) \mid \text{-}Ra \,\&\, \text{-}Ba \,\&\, S_ia)$ is negligible.

As per our derivation of the comparative result:

$P^{**}((x)(Rx \supset Bx) \mid \text{-}Ra \,\&\, \text{-}Ba \,\&\, S_ia) = P^{**}((x)((Rx \,\&\, S_1x) \supset Bx) \,\&\, \ldots \,\&(x)((Rx \,\&\, S_{i-1}x) \supset Bx) \,\&(x)((Rx \,\&\, S_{i+1}x) \supset Bx) \,\&\, \ldots \,\& (x)((Rx \,\&\, S_nx) \supset Bx) \mid (x)((Rx \,\&\, S_ix) \supset Bx) \,\&\, \text{-}Ra \,\&\, \text{-}Ba \,\&\, S_ia)$. $P^{**}((x)((Rx \,\&\, S_ix) \supset Bx) \mid \text{-}Ra \,\&\, \text{-}Ba \,\&\, S_ia)$.

And since the first term in the product must be less than 1, the confirmation afforded "all ravens are black" is (at least as) negligible.

Cite this article: Ward, Barry. 2022. "Resolving the Raven Paradox: Simple Random Sampling, Stratified Random Sampling, and Inference to Best Explanation." *Philosophy of Science* 89 (2):360–377. https://doi.org/10.1017/psa.2021.26

https://doi.org/10.1017/psa.2021.26 Published online by Cambridge University Press