Studies in History and Philosophy of Biological and Biomedical Sciences xxx (2017) 1–10

Contents lists available at ScienceDirect

Studies in History and Philosophy of Biological and Biomedical Sciences

journal homepage: www.elsevier.com/locate/shpsc

Galton, reversion and the quincunx: The rise of statistical explanation

André Ariew [a,*], Yasha Rohwer [b], Collin Rice [c]

[a] Department of Philosophy, University of Missouri, Columbia, MO 65211, United States
[b] Department of Philosophy, Oregon Institute of Technology, Klamath Falls, OR 97601, United States
[c] Department of Philosophy, Bryn Mawr College, Bryn Mawr, PA 19010, United States

ARTICLE INFO

Article history:
Received 27 February 2017
Received in revised form
14 August 2017
Available online xxx

Keywords:
Francis Galton
Quincunx
Statistics
Explanation
Normal distribution
Causation

ABSTRACT

Over the last six decades there has been a consistent trend in the philosophy literature to emphasize the role of causes in scientific explanation. The emphasis on causes even pervades discussions of non-causal explanations. For example, the concern of a recent paper by Marc Lange (2013b) is whether purported cases of statistical explanation are "really statistical" or really causal. Likewise, Michael Strevens (2011) argues that the main task of statistical idealizations is to distinguish between the causal factors that make a difference to the phenomenon to be explained and those that do not. But, the philosophy literature poorly reflects the history of the development of statistical explanation in the sciences. Francis Galton's (19th century) explanation for the laws of heredity is our case. Galton's statistical explanation was both innovative for his time and influential to our contemporary sciences. The key points to understanding Galton's statistical explanation for reversion is that it is autonomous from the real-world biological properties that make up an instance of reversion while still being approximately true of many real-world biological phenomena. Ours is an expanded discussion of ideas originated in Hacking (1990) and Sober (1980). We will articulate these features and compare our account with that of Lange and Strevens.

Published by Elsevier Ltd.

"The typical [statistical] laws are those which most nearly express what takes place in nature generally; they may never be exactly correct in any one case, but at the same time they will always be approximately true and always serviceable for explanation" (Galton, 1877, p. 17)

1. Current trends in the philosophy of scientific explanation are causal centric

As Michael Strevens (2012) reminds us, a great preponderance of philosophical work in the area of scientific explanation has been focused on the study of causal relations. The motivation for causal approaches can be traced back to issues with Carl Hempel and Paul Oppenheim's account (1948) according to which many scientific explanations involve deductions from premises stating a generalization and some initial conditions. But, as the well-known flagpole case demonstrates, Hempel and Oppenheim's deductive account fails to adequately account for the explanatory role causal facts play.

We can derive the height of the flagpole from the length of its shadow in the way required by Hempel and Oppenheim's account but the unpalatable result is the cause being explained by the effect. Therefore, there is a problem calling derivations of flagpole heights from shadows genuine cases of explanation. In contrast, deducing the shadow from the height of the flagpole is a genuine explanation because it cites causes to explain their effects.

According to Strevens, ever since the flagpole case, much of the literature in philosophy of scientific explanation has set its agenda accordingly to the following three questions:

1. What are causal relations?
2. How do deductive derivations and other semantic apparatus in science represent causal relations?
3. Besides the need to represent the causal production of the explanandum, what other norms govern the construction of scientific explanations?

This agenda has produced an intellectually unsatisfactory literature in the philosophy of scientific explanation in several regards. Here are three.

First, Strevens' agenda neglects important scientific advances that ought to guide philosophers' discussions about the norms of

* Corresponding author.
E-mail addresses: ariewa@missouri.edu (A. Ariew), Yasha.Rohwer@oit.edu (Y. Rohwer), crice3@brynmawr.edu (C. Rice).

https://doi.org/10.1016/j.shpsc.2017.08.001
1369-8486/Published by Elsevier Ltd.

scientific explanation. These important advances are typically neglected because they do not easily conform to popular philosophical accounts of good scientific explanation. Our case study is just one instance: Francis Galton's statistical explanation for what he called the "law of heredity": the processes of heredity maintain a fixed distribution of variation across generations. We would expect the explanation for a biological phenomenon to reference biological mechanisms and processes. But, Galton's explanation for this phenomenon is surprising because his explanation only makes reference to a mathematical result. The intergenerational stability of the distribution of variation is a deductive consequence of the distribution of variation of the previous generation. It matters not whether the ensemble is composed of coin tosses, shots on target, heights of soldiers, or biological characters, if the frequency of characters in the ensemble is normally distributed (under ordinary or equilibrium conditions) then it can be deduced that in the next generation there will be a normal distribution of the same mean and dispersion. Consequently, the exceptional characters will revert as a deductive consequence of the Normal distribution (Hacking, 1990, p. 183).

Galton's explanation was a first in biology. Galton's novel use of statistics to explain a real-world phenomenon led him to develop the statistical techniques of correlation, linear regression, and a variety of standards for goodness of fit between data and theory still used by scientists. According to historian Stephen Stigler, Galton's work "represents the most important step in perhaps the single major breakthrough in statistics in the last half of the nineteenth century" (Stigler, 1990, p. 281). Further, Galton's mathematical theory of inheritance was the basis for the 20th century synthesis of Darwinian natural selection and genetical theories of inheritance. It also laid the groundwork for the development of modern population and quantitative genetics (Fisher, 1953, p. 5). For such an important scientific advance, one would think that philosophers of science would pay more attention to it. But, while the overall number of citations referring to Galton's techniques exceeds that of the flagpole literature by orders of magnitude, the opposite is true within the philosophy of scientific explanation literature.[1]

A second problem with the tradition that stems from Hempel and the flagpole is that it has brought an over-emphasis on the role of causes in scientific explanation.[2] The current trend in the literature on non-causal explanation is concerned with a kind of demarcation exercise: t[illegible] to determine whether purported cases of explanation are genuinely non-causal or only appear to be non-causal but are in fact truly causal. Example[illegible]lude[illegible]other dividing 23 strawberries among three children[illegible]3a),[illegible]y it is impossible to walk through Königsberg and [illegible]each of the seven bridges once and only once (Pincock, 2007[illegible]w[illegible]xplain the periodicity of cidada life cycles (Baker, 2005[illegible]d e[illegible]he nature of the differences between genetic drift and natural selection in the modern theory of natural selection. (Lange, 2013b), (Walsh, Lewens, & Ariew, 2002). We believe the demarcation question is largely irrelevant to understanding the nature of Galton's statistical explanation for reversion and we shall criticize Marc Lange's (2013b) work accordingly.

The key points to understanding Galton's explanation for the law of heredity is that the statistical explanation is autonomous from the real-world biological properties that make up an instance of reversion while still being approximately true of many real-world biological phenomena. Galton's statement in the epigraph is a good expression of the view. Galton's reversion explanation is autonomous in the sense that the mathematical features of a statistical equation that Galton cites are sufficient to explain biological reversion, even though the statistical equation fails to accurately represent any real-world events that make up any particular instance of reversion.[3] The idealized statistical explanation of reversion is sufficient on the condition (or to the degree) that the freque[illegible]istribution of the trait in the population is *approximately* [illegible]al. What determines the degree to which a real world populati[illegible]pproximates a Normal distribution depends on minimal m[illegible]al requirements of the system, namely that the ensemble is the result of numerous randomized trials whereby the probability of the outcome of the one is independent of the probability of the outcome of any other. We discuss the details below. To help clarify our position we will contrast our account of the nature of Galton's reversion explanation with that of Marc Lange's account of what makes regression explanations "really statistical" as opposed to latently causal.

A third problem with the entrenched Hempelian tradition is [illegible]at there is too much emphasis on univocal and general accounts [illegible] scientific inquiry. We ought to recognize the possibility that there exist multiple adequate accounts of good scientific explanation. We will criticize Michael Strevens (2016) account of the role of idealizations in scientific explanations in this regard. His exclusive focus on causal-difference makers fails to account for the autonomy of statistical explanation. The point of Galton's use of the deductive features of a statistical law to explain regression is not, as Strevens has it, to highlight causal-difference makers. Rather, the point is that the statistical laws do all the explaining without the need to refer to any causal features of the ensemble.

We won't extrapolate from the historical case to formulate a universal account of all statistical explanations.[4] Rather we hope to learn some valuable lessons about best-case practices—in this case, Galton's explanation for reversion—in order to extract a norm of scientific explanation. We mean to reset the project of the philosophy of scientific explanation by freeing it from the tradition that automatically thinks about flagpoles and causal relations. Our historical case refers not to philosophers, like Hempel and Oppenheim, thinking generally about science, but a scientist inventing a unique way of explaining a natural phenomenon.[5] To us, Galton's explanation is exemplary: it represents a norm of scientific explanation whose nature is not revealed by answering the three questions Strevens lays out for us. Ultimately we want to understand how the deductive properties of an idealized representation of a biological phenomenon can be, in Galton's words, "approximately true and always serviceable for explanation" of a real-world biological process. To fully appreciate Galton's innovation it is important to note the state of statistical methodology at the time of Galton's writing. As you shall see, by articulating Galton's response to the pioneers of the statistical methodology we will have a deeper understanding of the aspects of Galton's explanation we are highlighting.

[1] Exceptions include: Hacking (1990), Sober (1980), Lipton (2009), Lange (2013a, 2013b), Gayon (1998), Ariew, Rice and Rohwer (2015), Radick (2011).

[2] Even Strevens' third question is dependent on an account of causation because in order to articulate the "other norms" we need to know what distinguishes them from the norm of representing causal production.

[3] We adopt Ian Hacking's (1990) concept of "autonomy" account for the nature of the Galton's innovative statistical explanation. See also Sober (1980) and Ariew et al. (2015).

[4] We consider other cases in science in Rice, Rohwer, Ariew (in prep.), "Explanatory Schema and the Process of Model Building".

[5] We certainly don't claim originality for this way of doing philosophy of science. The works of Nancy Cartwright are exemplary.

2. Galton on the statistical laws of heredity

The title of Galton, 1877 address to the Royal Institution of Great Britain is "Typical Laws of Heredity". His aim was to explain a peculiar large-scale ensemble effect of heredity.[6] In each generation we detect a great deal of individual variation of heritable characters. Individuals are tall and short, heavy and light, strong and weak, and the variations admit of every gradation in between. But, Galton claims, from the view of averages, all biological characters exhibit a puzzling order undetectable from measurements of individual differences: the proportions of various types is constant. And, despite the various comparisons between individuals and their offspring, heredity preserves the fixed distribution spread of characters over the course of generations on the condition that external circumstances remain constant. Galton writes that from the statistical point of view "uniformity prevails" and that "the processes of heredity are found to be so wonderfully balanced, and their equilibrium to be so stable, that they concur in maintaining a perfect statistical resemblance" (Galton, 1877, p. 3). Throughout the essay Galton provides a smattering of evidence of this "law of heredity", from the geological record to published statistical charts to the results of his own experiment with pea plants.

The law of heredity has negative consequences for natural selection. Ordinarily, tall, heavy, or strong couples tend to produce tall, heavy, and strong children. Likewise short, light, and weak, couples tend to produce short, light, and weak children. However, it is rare for children to maintain the identical features of their parents. In general, like does not beget like. Nevertheless, surprisingly, tall children are rarely taller than the tallest parent, and, over the course of generations, tallness does not tend to maintain itself in the lineage. The extreme pairings does not affect the stability of the distribution pattern that the law of heredity describes. The frequency or measurement of either the extremely tall or the extremely short does not increase. Rather, the extremes appear to "revert" to the population mean. This is problem for any process of selection. Pigeon breeders selecting for an extreme shade of grey among their stock hope to eventually produce a generation where the desired character is more common. But, the reversion effect will thwart their progress.

The stability of the law of heredity also confounds selection for mean characters, not just the extremes. Galton illustrates with a hypothetical example of a race containing 100 giants and 100 medium-sized men. Considering the known hardships befalling the life and reproductive capacities of giants, we would expect the 100 medium men to be more fertile, more likely to breed with like kind, and better fitted to survive hardships, &c". Giants are known to suffer from a consumptive constitution and "languid" circulation. Their kind is also likely to be diluted by marriage. For reasons such as these, we should expect there would be fewer giants and more medium-sized men in the second generation than in the first. But, according to the law of heredity, the proportion of giants to medium-sized men in the first generation will remain the same in the second generation. Galton concludes: "The question, then, is this: How is it, that although each individual does not as a rule leave his like behind him, yet successive generations resemble each other with great exactitude in all their general features?" (Galton, 1877, p. 2).

As historian Stephen Stigler points out, Galton's presentation is a more articulate version of a problem that Fleeming Jenkin recognized in his 1867 review of the Origin of Species (Stigler, 2010). The medium-sized men in the thought experiment have heritable advantages over the giants in the struggle for life. Yet, the balancing forces of heredity prevent selection from fundamentally altering the stable frequency of variation. In the Origin of Species, Darwin claimed that the effects of "reversion" was an exaggeration and natural selection would eventually overcome any such barriers. It would turn out that Darwin was right but it took 20th century evolutionists, like R.A. Fisher to show how he did so, within the framework of Galton's explanation for the stability of heredity. Galton's aim in the 1877 essay was not to solve the problem for Darwin but to explain why the distribution of characters is stable over the course of generations?

2.1. The law of deviation and Quetelet's chart of heights

Galton declares that we gain an insight into the puzzle of heredity when we bring to bear Adolphe Quetelet's discovery that the shape of the constant distribution of characters tends to conform to a precise mathematical law: "the amount and frequency of deviation from the average among members of the same race, in respect to each and every characteristic, tends to conform to the mathematical law of deviation." (Galton, 1877, p. 2). The law of deviation is Galton's term for what was otherwise known as the "law of errors", first introduced in the 18th century by the mathematician Abraham de Moivre as a graphical expression of the outcome of coin tossing trials. Toss a coin n-times and indicate the proportion of heads to the total number of tosses on the x-axis of a graph from 0 to n. The y-axis represents the number of times in the sequence of trials where the coin lands heads. As the number of tosses increases, n gets larger (without bound), the resulting graph increasingly resembles a bell-shaped curve with its peak at the mean value and sloping sides representing the amount of dispersion around the mean (see Hacking, 1990).[7] Ever since de Moivre, natural philosophers believed that the statistical stability—expressed as a mathematical equation—was a sign of divine intervention, and, as such, had wider application. In 1844, Quetelet introduced the idea that most, if not all, human attributes, from height to moral attitudes, conform to the law of error. Galton goes one step further by extending the scope of the law to all biological characters and changes the name to the "law of deviation". The "curve of error" becomes, in Galton's later writing, the "normal distribution". The name changes are relevant when we discuss the differences between Quetelet and Galton's methodology in a later section of this essay.

To illustrate how human heights approximate the normal distribution, Galton reproduces Quetelet's statistical table on heights of American, French, and Belgian soldiers (see Fig. 1). The data is arrayed in ascending order from tallest to shortest. In each series there are two columns titled "observed" and "calculated". The numbers therein represent the frequency of individuals that correspond to the unit measured. "Calculated" refers to frequencies

[6] Galton's statistical explanation for the "laws of heredity" is not Galton's only concern. He had theories about the mechanisms for transmission as well. He adopted a gemmule theory in which hereditary material is transmitted by the blood. See Radick (2011) for a good discussion. However, as Hacking (1990) points out, Galton recognized his statistical explanation as a distinctive achievement quite apart from his mechanistic theories of heredity. The account we provide is not affected at all by the specifics of Galton's theories of transmission beyond the justification of his statistical assumptions. That's because the *explanandum* is how the distribution of statistical variation is maintained over generations, not how individual differences aggregate to produce a variational pattern.

[7] Today we would say that de Moivre identified the conditions for the binomial distribution, a special case of the central limit theorem developed by LaPlace and Gauss. The central limit theorem has much wider application because it appears in different scales and different domains, from processes that aggregate fluctuations from an average value. The aggregates can be additive, multiplicative, even distributions on the log scale (McElreath, 2015, p. 78, p. 78).

Scale of Heights.	American Soldiers (25,878 Observations).		France (Hargenvilliers).		Belgium, Quetelet. 20 years' Observations.	
	Observed.	Calculated.	Observed.	Calculated.	Observed.	Calculated.
mètres.						
1·90	1	3				
1·90	7	5	..	..	1	1
·87	14	13	..	1		
·84	25	28		3	2	3
·81	45	52	25	7	7	7
·79	99	84		16	14	14
·76	112	117	32	32	34	28
·73	138	142	55	55	48	53
·70	148	150	88	87	102	107
·68	137	137	114	118	138	136
·65	93	109	144	140	129	150
·62	109	75	140	145	162	150
·60	49	45	116	132	106	136
·57	14	24	..	105	110	107
·54	8	11		73		53
·51	1	4		44		28
·48	..	1		24		14
·45	..	..	286	11	147	7
·42	..	..		4		3
·39	..	..		2		1
·36	..	..		1		
	1000	1000	1000	1000	1000	1000

Degrees of Dynamometer.	Lifting Power of Belgian Men.	
	Observed.	Calculated.
200	1	1
190	9	10
180		
170	23	23
160		
150	32	32
140		
130	22	23
120		
110	12	10
100		
90	1	1
	100	100

Fig. 1. Aldolphe Quetelet's published chart on the height of American, French, and Belgian soldiers, reproduced in Galton (1877).

of an abstract population that approximates the mathematical law of error. The "striking" fact that Galton wants us to see from the charts is that despite differences between the three nationalities, in all instances, the observed height distribution closely resembles or approximates the calculated distribution from the law of deviation, meaning that height and strength distribution for each population more or less conforms to the normal distribution just as the law of deviation dictates.

To Quetelet the significance of the chart of heights is that each nationality has its own signature "curve of error", each with distinct means. This is the idea behind Quetelet's famous concept of the "average man".) Quetelet cites this as evidence that there are essential developmental differences between Americans, French, and Belgians. But, Galton's explanatory aims are distinct. He intends to explain the universality of the law of deviation in biological characters, both across disparate populations and over the course of generations within a reproducing population. He writes: "although characteristics of plants and animals conform to the law, the reason of their doing so is as yet totally unexplained." That's why the different height means don't matter to Galton, as they do to Quetelet.

Why does conformity of the data to the normal distribution matter to Galton? Galton's aim, his method for explaining the law of heredity is to turn a biological problem into a mathematical problem. Galton states his aims clearly:

> "The outline of my problem of this evening is, that since the characteristics of all plants and animals tend to conform to the law of deviation, let us suppose a typical case, in which the conformity shall be exact, and which shall admit of discussion as a mathematical problem, and find what the laws of heredity must then be to enable successive generations to maintain statistical identity." (Galton, 1877, p. 4)

Please cite this article in press as: Ariew, A., et al., Galton, reversion and the quincunx: The rise of statistical explanation, Studies in History and Philosophy of Biological and Biomedical Sciences (2017), https://doi.org/10.1016/j.shpsc.2017.08.001

But, before Galton can treat the puzzle as a mathematical problem he has to justify why he is doing so. He has to explain how is it possible for a real-world ensemble of characters to obey the law of deviation. Strictly speaking, the law of deviation is an idealization. No real-world ensemble of characters is normally distributed. The bell-shaped curve it generates reflects the parameters of a mathematical equation. As we see in de Moivre's coin tosses and Quetelet's chart of soldier heights, real-world ensembles can only tend to conform to the normal distribution as the number of independent samples increase (without limit). But, coin tosses and the biological processes involved in determining soldier heights are physically disparate. How do such disparate systems conform to a single statistical distribution curve?

The answer is that the requirements for a real-world population to approximate the normal distribution are minimal. We call these the "minimum material conditions": There needs to be a process where the events are randomly sampled. The sample size needs to be relatively large. And, the sampling outcomes must be independent, meaning that the outcome of one event or trial has no effect on the outcome of any other. Sample size matters because approximation or conformity is a matter of degree. While the sample of heights show rough conformity, we can presume that a larger sampling of heights will conform to an even greater degree. Galton assumes that even the rough conformity justifies modeling solider heights with the idealized law of deviation.

To demonstrate how a real-world process that features the minimum material conditions aggregates to approximate the normal distribution, Galton utilized a "quincunx", a mechanical apparatus that he devised for the address.

The quincunx is a shot-dropping machine with a glass face, resembling a Japanese pachinko machine (see Fig. 2). At the top is a funnel from which to drop small pellets. Upon passing the narrow neck of the funnel the pellets cascade across an array of 17 rows of offset spikes. The pattern of spikes is where the name "quincunx" comes from—it resembles the mesh pattern like those of fishermen's nets. At the bottom of the machine is a row of compartments separated by a thin wall. This is the final resting place for the dropped pellets. Galton writes: "I will pour the pellets … from any … point above the spikes; they will fall against the spikes, tumble about among them, and after pursuing devious paths, each will finally sink to rest in the compartment that lies beneath … " (Galton, 1877, p. 5). When encountering each of the numerous rows of spikes, the pellets fall either right or left. In the commonest case a pellet falls to the right and left with equal frequency and comes to rest at center— directly below the point at which it was dropped. Less commonly a pellet falls to one side more frequently than to another and its final resting spot is some distance to the left or right of center. The pellets resting at the greatest distance from center are rarest, because they require a long run of falling to the same side of each pin they encounter. The result of all these runs is a pile of pellets whose outline resembles the normal distribution.

Fig. 2. Galton's that still exists in University College, London.

The quincunx simulation shows what de Moivre showed with coin tosses could be applied to heights, or, presumably other hereditary characters. The law of deviation is widely applicable because, although it is never precisely replicated in any real-world finite population, many real-world ensembles feature the minimal material conditions to generate a bell-shaped distribution pattern. Galton writes: "the law of deviation is purely numerical; it does not regard the fact whether the objects treated of are pellets in an apparatus like this, or shots at a target, or games of chance, or any other of the numerous groups of occurrences to which it is or may be applied" (Galton, 1877, p. 7).

Galton can now turn a puzzle about biological heredity into a mathematical problem. The original question, why do processes of heredity concur to "maintain a stable statistical resemblance"? can be restated: how is the normal distribution maintained over the course of generations? Galton will eventually answer with a modified version of his quincunx. But, first, he must show the law of heredity applied in a biological case. Neither Quetelet's chart of heights nor the first quincunx demonstration has a hereditary component. To remedy, Galton reports on a breeding experiment on sweet peas that he conducted some time earlier.

2.2. Galton's sweet peas

Galton chose the sweet pea for his experiments after consulting with Darwin. The peas are hardy, convenient, and, most importantly, they don't cross-fertilize—a confounding factor in the study of heredity. First he arrayed thousands of seeds according to the frequency distribution of weight. Out of that lot, he selected nine sets, each containing seven packets with ten seeds of identical weight. The first packet contained the giant seeds, +3 degrees of deviation, the seventh the very small seeds, −3 degrees of deviation, and the rest contained the intermediates, each marked by a distinct degree of deviation from the mean of the original population of thousands. He sent the sets to a variety of friends and acquaintances in the UK, Darwin included, instructing each of them to plant and report the produce in terms of the distribution of seed weight among the offspring. We do not know which packet Darwin possessed. Two sets were failures but among the seven packets returned (containing 490 carefully weighed seeds in total) the final result conformed to the law of deviation. had it been the case that heredity follows the simple rule, like begets like, the extreme sizes would have reshaped the whole distribution curve towards more extremes. But, that is not what the experiment yielded. The seed weight within each packet was Normally distributed and equally dispersed. And each was distributed about a value that was closer to the average population weight than the average of their parents. The overall population, including the extreme groups, reverted to the original mean of the overall parental population. The sweet pea experiment demonstrated that the processes of heredity conspired

Please cite this article in press as: Ariew, A., et al., Galton, reversion and the quincunx: The rise of statistical explanation, Studies in History and Philosophy of Biological and Biomedical Sciences (2017), https://doi.org/10.1016/j.shpsc.2017.08.001

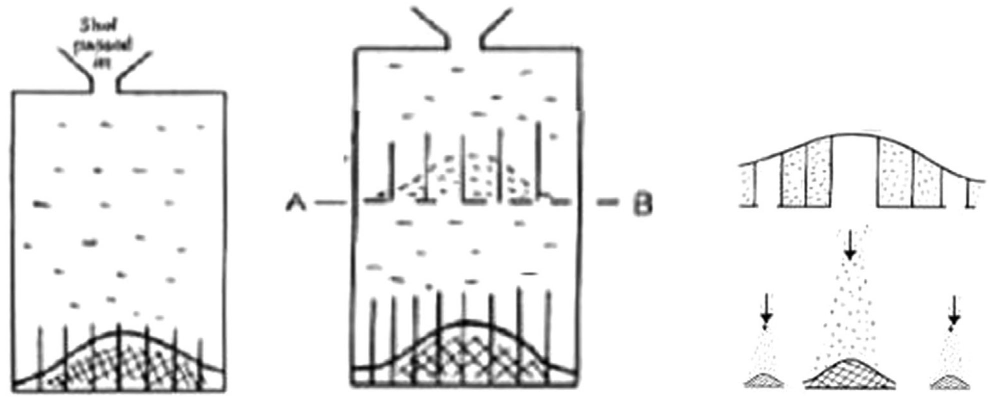

Fig. 3. A drawing from Pearson's biography based on a letter from Galton to George Darwin, reproduced in Stigler (1986, p. 278).

to maintain an approximately normal distribution of seed characteristics over two generations.

In sweet peas, the law of heredity holds; i.e. the frequency distribution of traits is stable across generations. Now, we're back to the original question: why does the law of heredity hold for biological heredity? Recall, the first quincunx didn't answer that question. The first quincunx demonstrated why Galton's statistical assumption—that the distribution of characters approximates the normal distribution—is justified. It is justified because the world is full of approximately normal distributions. That is, the world is full of populations whose characters we can treat as numerous random trials with independent outcomes. The normal distribution describes the aggregation of these sample trials. As Galton tells us at the outset, his methodology to explain why biological heredity maintains the normal distribution over the course of generations is to assume a case where the conformity to the normal distribution is exact, "which shall admit of discussion as a mathematical problem" (Galton, 1877, p. 4). Ironically, Galton's address contains no math (except in an appendix). Rather, he simulates the mathematical results with a thought experiment involving a modified quincunx.

2.3. Deduction and quincunx 2.0

Galton's virtual quincunx contains a second field of pins beneath the compartments of the original quincunx (see Fig. 3). [8] The individual compartments are also modified; each has a trap door underneath them. ~~So, rather than starting from the same funnel point~~, each pellet begins their run from the compartment that they settled into during the first run. The pellets in the right-most compartments represent the heaviest type; the lightest are in the left-most compartments. These pellets are then dropped from each of the trap doors, bounce around the spikes and settle on the bottom, each forming a normally distributed heap of pellets. When all the groups are merged into one, the shot again arrays itself into a Normal distribution composed of the aggregate of the distributions of the subsets. This "grand normal" distribution is just the aggregate of the normal distributions found in each of the subgroups. As Galton explains, "Heap adds itself to heap, and when all the pellets have fallen through, we see that the aggregate contributions bear an exact resemblance to the heap from which we originally started" (Galton, 1877, p. 8).

This is the same result seen in the sweet pea breeding experiment. When taken together, the offspring of all the individual sets of cultivated seeds were normally distributed around the mean, not of each packet, but of the entire parental population. Likewise, the offspring of parents from each category of size were normally distributed. The sweet pea experiment acted exactly in the way that the quincunx predicts. Galton writes: "The conclusion is … that the processes of heredity must work harmoniously with the law of deviation, and be themselves in some sense conformable to it."

In a later book, *Natural Inheritance*,[9] Galton explains reversion with his quincunx (see Fig. 4)[10] When pellets in the top half of the quincunx are released their average end point is directly below. But, what of pellets resting in a lower-level compartment: where do they come from? The answer to the second question is not "directly above". Rather, on average, the pellets at the bottom of the quincunx come from a compartment towards the middle of the previous generation. This is because in a normally distributed ensemble, the number of pellets gets larger as we count from the outermost compartments inwards towards the center. So, there are more pellets that can wander from the center towards the extremes than there are pellets that can wander from the extreme towards the center. In his 1889 book Galton no longer calls the phenomenon "reversion", which traditionally indicated an empirical phenomenon well known to Darwin's contemporaries. Rather, Galton calls it "regression", ~~which is a purely statistical phenomenon.~~ Historian Stigler concludes: "Galton's great insight from this new approach was that stability implied … regression. [T]he entire puzzle was resolved by this one fundamental insight" (Stigler, 2010, p. 477). The insight is that intergenerational stability could be explained by reference to the mathematical properties of the law of deviation, modeled by the quincunx which instantiates the minimal material conditions required for the law of deviation to hold.

3. Galton's statistically autonomous explanation vs. Quetelet's statistical reductive explanation

According to Ian Hacking (Hacking, 1990), Galton's explanation for the law of heredity and reversion was unique in science. It was the first instance of what Hacking calls "statistically autonomous" explanation. Autonomy stands in contrast to reduction but without commitment to metaphysical antireductionism. Galton was not arguing that any one instance of the heredity law in action was irreducible to underlying deterministic principles. Rather he was explaining a biological phenomenon while leaving out the

[8] Reproduced from (Stigler, 2010). The drawing comes from Pearson's biography based on a letter from Galton to George Darwin.

[9] At this point Galton has already discovered the analytic powers of regression. See Stigler (2010).

[10] Taken from Stigler (2010). Stigler drew in the arrows.

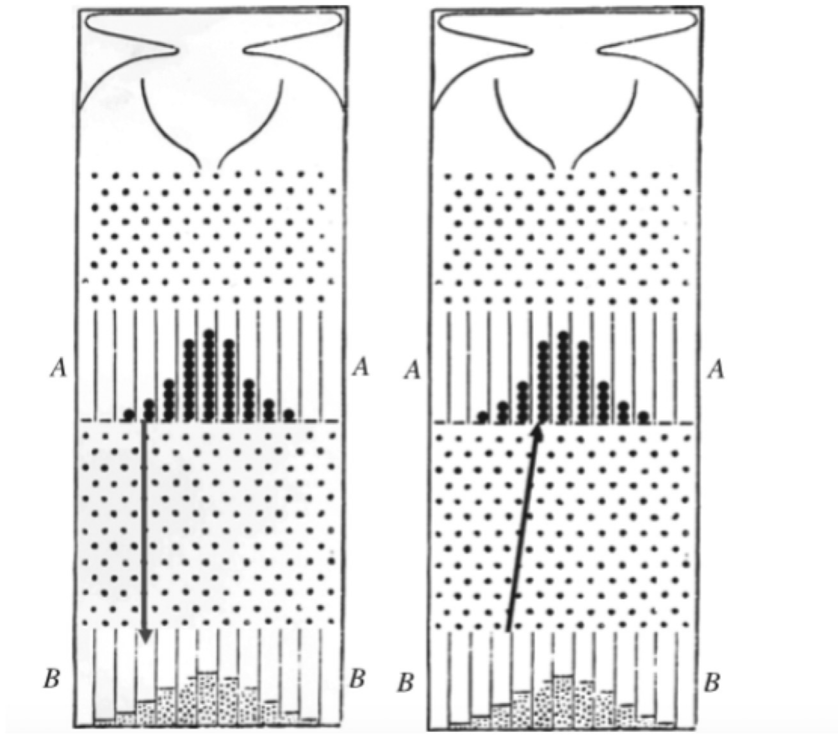

Fig. 4. An illustration of the regression phenomenon, reproduced in Stigler (2010).

biological details. Galton attributes to Quetelet the central statistical insight for his explanation of the law of heredity. But, it was Galton who invented the statistically autonomous explanation for the law of heredity. Quetelet did not. That is because between Galton and Quetelet, the distribution patterns and conditions that generated each graphical curve were the same but the biological interpretation was distinct (Ariew, Rice, & Rohwer, 2015). The aim of Quetelet's "social physics" was to identify the essential differences between races and populations, as exemplified by the chart of soldier heights that Galton showed his 1877 audience. The different means were relevant for Quetelet's inquiry, the dispersion was "error", or more precisely, the reflection of "accidental causes" of development. Hence, following de Moivre, the statistical law was the "law of error" and its graphical representation the "curve of error".

Galton's project, however, was to explain the maintenance of the distribution (same means, same dispersion) over the course of generations. Both the means and the tail ends of the distribution curve matter for Galton. Hence Galton renamed the statistical law "the law of deviation" and its graphical representation the "normal distribution".

In order to see the differences between these two approaches to applying statistical modeling techniques, let's look a little more closely at Quetelet's explanatory project. Quetelet's project provides the historical context to see what made Galton's statistical explanation unique and why Hacking is right to contrast Galton's "statistically autonomous" explanation from reductive explanations.[11]

Quetelet was an astronomer turned social scientist. From observational astronomy he borrowed the idea—pioneered by LaPlace and Gauss—that the appearance of the bell-shaped distribution pattern for observational data reveals the difference between "error" and "truth". The pressing problem in observational astronomy was to determine the trajectory of a planet. A technician had to rely on the various observations made by different astronomers at different times and places with a variety of observational techniques, taking into account that many reports were fraught with error. LaPlace and Gauss's insight was to attend to the aggregate of the observations, not any subset. The larger the number of observations, the more likely the observational errors will, as it were, get swamped by the preponderance of data, which tends to conform to the planet's true trajectory (assuming the observations meet the minimal material conditions of being numerous and independent). Quetelet defined his "fundamental principle" of social physics accordingly: "*the greater the number of individuals observed, the more do individual peculiarities, whether physical or moral, become effaced, and leave in a prominent point of view the general facts, by virtue of which society exists and is preserved*" (Quetelet, 1835, p. 6, his italics). But, rather than distinguishing "error" from "truth", Quetelet was distinguishing between "constant" and "accidental" causes of human development. The average or mean value reveals information about the "constant" developmental causes of height: if the constant causes of development were acting alone, every individual Belgian soldier would attain the same height.[12] Dispersion around the mean is the effect of the various external "disturbing" causes preventing each Belgian soldier from attaining the true height of its race. Differences between the Belgian, French, and American averages is evidence of the differences in developmental causes between the groups.

Now we see why Quetelet called the bell-shaped distribution pattern the "curve of error". The disturbing causes create error. They misrepresent the normal state of the population. As Elliott Sober put it, "For Quetelet, the point of attending to variation was to *see through it*—to render it transparent. Averages were the very

[11] Margaret Morrison (2014) has an excellent discussion of "irreducibly statistical phenomenon" that differs in some ways from our account.

[12] Just as if every pellet that fell from its funnel experienced no interfering pin, each would hit the bottom of the quincunx in exactly the same place.

antithesis of artefacts; they alone were the true objects of inquiry" (Sober, 1980, p. 36).

Quetelet's analysis of population averages had broad and important implications for the burgeoning field of sociology. He developed an early form of analysis of correlation[13] to explain a puzzling phenomenon first discovered in the 17th century and persistent in the census tabulations for every region in the world. Despite the variation in households of numbers of male and female births, the aggregate yields a stable sex ratio skew towards male births. The consensus explanation (from DeMoivre, among others) was divine intervention. But, Quetelet applied his analysis of averages across various tabulations to come up with a hypothetical cause. The method involved comparing the sex ratios among various sub-populations against the overall average and seeing which category of tabulation made the difference. Averages were unaffected by a host of comparisons: urban vs. rural births, warm vs. cool climates, higher vs. lower elevations. But, he found a difference in comparing births from married couples and from "illegitimate" arrangements. The excess of males to females that is indicative of the entire European population (5%) is the same as the excess among children from married couples and more than the excess among illegitimate children (3%). Assuming married couples' males tended to be older than females, Quetelet concluded that the relative ages of the parents physiologically influence the sex of the offspring. This is a good example of Quetelet's reductionistic statistical methodology that seeks to use averages to uncover the underlying causes of the population's distribution. For Quetelet, only the averages in a curve of error matter because they allow us to discover the true causes of development.[14]

Galton, however, was aware that Quetelet's social physics could not explain the law of heredity. Galton writes:

> "let me point out a fact which Quetelet and all writers who have followed in his path have unaccountably overlooked, and which has an intimate bearing on our work tonight. It is that, although characteristics of plants and animals conform to the law, the reason of their doing so is as yet totally unexplained" (Galton, 1877, p. 8).

You don't sufficiently explain the intergenerational stability of the normal distribution by stating that the normal distribution is the product of a host of a lot of independent causes. As Galton puts it "the processes of heredity that limit the number of the children of one class, such as giants, that diminish their resemblance to their fathers … are not petty influences, but very important ones" (p. 8).

The answer Galton is seeking is found in the statistical consequences of *any* real-world population (from coins to pellets in the quincunx to biological hereditary processes) that approximately conforms to the idealized law of deviation. Conformity to the idealized law of deviation is the consequence of the aggregate of the host of independent causes, as the first quincunx demonstrates. But, the explanation of the law of heredity is not found by attending to these instances.

In reflecting upon the kind of explanation Galton has provided, he writes the account description found on the epigraph of this essay. Hacking cites this passage as evidence that Galton was aware of the novelty of his explanation (Hacking (1990), p. 180). Sober (1980) concurs.

3.1. *Galton's explanatory methodology: a summary*

In our presentation we described Galton's explanatory methodology along the way. It would be useful now to lay it out before comparing Galton's account of statistical explanation with that of Lange and Strevens.

Galton's *explanandum* is the law of heredity: the intergenerational maintenance of variation over the course of generations. He is asking, Why do the processes of heredity concur to maintain statistical resemblance? Galton's explanatory strategy is to begin with a *statistical assumption*: hereditary characters *approximate* the normal distribution, the graphical representation of the law of deviation. The purpose of the assumption is to treat the biological phenomenon as a mathematical problem.[15] With the first quincunx he provides justification for his statistical assumption. Nature is full of processes that generate a normal distribution because the *material requirements are so minimal*: the character samples resemble trials of numerous independent events.[16] The phenomenon to be explained is a *mathematical consequence* of the law of deviation. Galton demonstrates this with modifications of the first quincunx. If a population of characters is approximately normally distributed then it can be *deduced* that in the second generation there will be a normal distribution of about the same mean and dispersion, assuming that the external conditions remain constant. A consequence of this deduction is the phenomenon of reversion. The exceptional members of the offspring generation will typically not be descendant from the exceptional members of the parent generation. (Hacking, 1990, p. 186). This is Galton's mathematical *interpretation* of the laws that the biological processes must concur. Galton's explanation is statistically autonomous because he is explaining a biological phenomenon without referring to any of the causes underlying any particular instances of the phenomenon. He interprets the deductive features of the normal distribution as "approximately true" of any instance of the law of heredity—never exactly true—but due its generality and approximation, "always serviceable for explanation."[17]

4. Galton versus Lange and Strevens on statistical explanation

Let's compare our analysis of the explanatory structure of Galton's explanation of the laws of heredity with that of two of our contemporaries, Marc Lange and Michael Strevens. Both are in some way or other committed to the Hempelian program we discussed at the outset where inquiry into scientific explanation is focused on questions concerning causation. Lange is concerned with demarcating between causal and non-causal, statistical explanations. Strevens is concerned with the role that causal factors play in idealized statistical explanation.

4.1. *Lange's "really statistical explanation"*

In a recent paper, Marc Lange (Lange, 2013b) argues that regression explanations are a distinct kind of statistical explanation. They are "really statistical" because they show the result to be merely a "statistical fact of life". In contrast, some statistical

[13] R.A. Fisher (1953) acknowledges that Quetelet's is a precursor to his own "analysis of variance".

[14] It is worth noting that Darwin referenced Quetelet's conclusion to his study on ratio skew in a notebook entry dated just after he read Malthus on populations. For a detailed discussion see (Ariew, 2007).

[15] See also Ariew et al. (2015).

[16] Aidan Lyon (2014) provides compelling reasons why there is more to the story about why normal distributions are normal.

[17] Stephen Stigler (2010, p. 478–479) notes that in 1889, philosopher and psychologist, Hiram M. Stanley correctly pointed out that heredity and environment were inextricably confounded in Galton's data, hence, rendering his model of heredity tenuous. While Galton never replied to Stanley, R.A. Fisher successfully utilized Galton's theoretical model to construct the fundamental mathematical basis for Mendelian genetics.

Please cite this article in press as: Ariew, A., et al., Galton, reversion and the quincunx: The rise of statistical explanation, Studies in History and Philosophy of Biological and Biomedical Sciences (2017), https://doi.org/10.1016/j.shpsc.2017.08.001

explanations are causal because they invoke relevant features of the event's causal history. To illustrate the difference, Lange describes the results of a fair coin tossed 100,000 times (p. 171). Bracket out runs of 20 consecutive tosses beginning with toss numbers that differ by 10, e.g. 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. But, runs with exceptionally high number of heads tend to be followed by runs of fewer numbers of heads simply because the probability of such a high number of heads is always low compared to the other possible outcomes. The causal explanation invokes the relevant features of the coin and tossing mechanism that determine its fair propensity (assuming that each toss outcome is independent of the others). The chance that a run with an exceptionally high number of heads will be followed by a run with fewer heads is high. It is just the chance that at least one tails will appear in some number of tosses following the exceptional high number of heads. Lange tells us that for 10 tosses, the chance that at least one tails appears is 1 − (0.5) 10 = 1023/1024. According to Lange, this explanation is causal because "it explains by virtue of describing relevant features of the result's causal history" (p. 172).

An alternative explanation invokes regression to the mean. Accordingly, it suffices to say that we should expect runs of heads to follow from runs of tails as "fallout from the statistical character of the case". There is no need to invoke the coin's propensities or facts about the tossing mechanism because, according to Lange, the explanation is not "deepened" by these physical facts. A regression explanation states the mere fact "there is a statistical association between the outcomes" of overlapping coin toss runs. In general, what's required of Lange's "really statistical" explanations is only that they identify a "particular signature of statistical processes that the explanandum exemplifies" (p. 177).

Lange detects a bright line between the causal and statistical explanation because he is leaving out relevant details for each account. If we put back in the details of the two explanations, the two cases appear to be two ways of providing the same kind of explanation.

Let's start with the "causal" explanation. Lange claims that the relevant features are the causal historical details about the coin tossing set up. But, take Galton's explanatory strategy as a model and you see that what is signifi[illegible]t the coins are not the forces of the tossing mechanism o[illegible] trajectories of each coin in the actual experiment. Rather, the relevant features are those minimum material conditions about the coin tossing set up that allow us to treat the puzzle about the coins as a mathematical problem from which Lange provides his calculations. The relevant causal information establishes that there are two equiprobable outcomes for each trial (heads or tails with a fair coin), the tosses are independent, and numerous (100,000). Violate any one of these conditions and Lange's calculations from an idealized equation would appear to be inapplicable to the real-world case. Further, Lange assumes that the actual coin flips approximate the conditions require to apply the standard axioms of probability theory, as first described by Russian mathematician A.N. Kolmogorov in 1933. Otherwise, his calculation, for 10 tosses, the chance that at least one tails appears is 1 − (0.5)10 = 1023/1024, is not applicable to the chances of a coin.

Turning to Lange's "really statistical explanation", without the statistical assumption that the coin tosses conform to the law of large numbers (among others), references to a "particular signature of statistical processes that the explanandum exemplifies" has no real-world referent because all real-world cases can only *approximate* the statistical equations used in the explanation. Put differently, this is not an explanation of the coin tosses unless the coin tossing set up is shown to approximate the idealized law of regression. But this is just to say that the minimal material conditions are also required for the really statistical explanation.

Lange ig[illegible] the roles of what we've been calling "approximation" and [illegible]imal material conditions" in the background assumptions of both his causal and really statistical explanation for coin tossing regression. Perhaps the oversight is due to his focus on his prior commitment to the kind of causal-centric accounts of scientific explanation that we described at the outset.

4.2. *Strevens' kairetic account*

Michael Strevens' account of explanations for large-scale effects is encapsulated in his "Kairetic account" (Strevens, 2011). But, more recently, he offers a more general account of the explanatory power of idealized explanations, explanations that involve some degree of distortion of reality (Strevens, 2016). Accordingly, "the role of idealization … is to indicate that certain factors make no difference to the phenomenon to be explained" (p. 2). And, by "certain factors", and "difference making", Strevens is referring to causes.

A virtue of Strevens' account is that it applies to both explanations of individual events and more abstractly to ensemble regularities. An explanation for the extinction of a particular species, for example, requires that "only aspects of the causal history that made a difference to whether or not the event occurred earn a place in an explanatory model" (p. 4). Other factors causally contribute to how the extinction occurred, like the fact that individuals living in a particular region died out first or the exact date of the final death. But, on Strevens' account, these should not be part of an idealized explanation unless they also made a difference to the fact that the extinction occurred.

An explanation for an abstract regularity such as gases' conforming to Boyle's law also follows the same causal abstraction procedure, according to Strevens. An explanation "describes just those aspects of the causal process that make a difference to the fact that a gas follows the curve". Causal facts about the gas that merely affect how the gases follow the curve in an individual case are irrelevant to the explanation. So, a good explanatory model for the Boylean patterns omits all those factors particular to any individual instance of conforming to Boyle's law and "replacing them with an abstract specification of the gas's properties that is *instantiated* both by gases that have collisions and those that do not, by gases in which there are long-range attractive forces between molecules and those in which there are not, and so on" (p. 4).

Strevens' account falters when he insists on the generality of his difference making explanation procedure: "Idealizations across the sciences should be interpreted in the same way" (p. 5). From the outset we warn against over generalizing from a philosophical account of scientific explanation or idealization.[18] In this case, Strevens fails to recognize Galt[illegible] statistically autonomous explanation for the laws of heredity[illegible]ton's explanatory strategy is not as Strevens' describes. Galton's use of an idealized statistical model (or equation) is not about distinguishing between genuine causal difference makers and non-difference makers of any particular features of biological heredity. Rather, the strategy is to treat the biological phenomenon to be explained as a mathematical problem. To this end, Galton demonstrates with the quincunx that the idealized mathematical law of deviation applies to a wide range of real-world phenomenon because many real world processes feature the minimum material conditions. It is true, as Strevens says, that the quincunx describes just those aspects of the causal

[18] Another issue is that Strevens's account seems to leave little [illegible]for a clear distinction between the roles of idealization and abstraction (~~[illegible] for peer review~~).

Please cite this article in press as: Ariew, A., et al., Galton, reversion and the quincunx: The rise of statistical explanation, Studies in History and Philosophy of Biological and Biomedical Sciences (2017), https://doi.org/10.1016/j.shpsc.2017.08.001

process that make the difference to the fact that the aggregate of the pellets approximates a normal distribution pattern. It is also true that this is one part of Galton's explanatory strategy, the part where he justifies his statistical assumption that biological characters approximate the normal distribution. But, the reasons they do so is not Galton's ultimate aim. His aim is to explain the stability of intergenerational variation. Distinguishing difference makers from non-differences makers no more explains the than Quetelet distinction between constant and accidental causes of character development (Ariew et al., 2015).

Contrary to Strevens, the idealizations used in explanations across the sciences should not all be interpreted as distinguishing causal and non-causal difference makers. Some do while some, like Galton's use of the normal distribution, do not.

5. Conclusion

The causal agenda set ever since the flagpole counter-example to Hempel's account of scientific explanation fails to appropriately account for the nature of Galton's influential and important explanation for the law of heredity. Galton's statistical explanation is statistically autonomous: it explains reversion without regard to the causes that actually underlie any individual case. That does not mean that Galton believed there are no causal explanations for any of the instances. Rather, it means that the causes are irrelevant to the explanation of reversion. Reflecting upon the three questions that Strevens articulates for a Hempelian-post-flagpole agenda, we aver that Galton's exemplarly explanatory strategy is best articulated by ignoring these questions altogether.

Acknowledgment

We would like to thank the following for excellent feedback on earlier drafts: Lynn Chiu, Philippe Huneman, Marc Lange, Charles Pence, Michael Strevens, Denis Walsh, an anonymous historian who reviewed the article for this journal, and audiences at the following venues: Department of Philosophy at the University of Connecticut, ImmunoConcept Lab at the University of Bordeaux, the Department of Philosophy at the University of Kansas, and the workshop, "Unifying the debates: mathematical and non-causal explanations" at the IHPST in Paris.

References

Ariew, A. (2007). Under the influence of Malthus's law of population growth: Darwin eschews the statistical techniques of Aldolphe Quetelet. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 38*, 1–19. https://doi.org/10.1016/j.shpsc.2006.12.002.

Ariew, A., Rice, C., & Rohwer, Y. (2015). Autonomous-statistical explanations and natural selection. *The British Journal for the Philosophy of Science, 66*(3), 635–658.

Baker, A. (2005). Are there genuine mathematical explanations of physical phenomena? *Mind, 114*, 223–238.

Fisher, R. (1953). The expansion of statistics. *Journal of the Royal Statistical Society. Series A (General), 116*, 1–10.

Galton, F. (1877). Typical laws of heredity. *Proceedings of the Royal Institute of Great Britain, 8*, 282–301.

Gayon, J. (1998). In *Darwinism's struggle for survival: Heredity and the hypothesis of natural selection* (n edition). Cambridge, U.K.; New York: Cambridge University Press.

Hacking, I. (1990). *The Taming of Chance* (1 edition). Cambridge England; New York: Cambridge University Press.

Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science, 15*, 135–175.

Lange, M. (2013a). What makes a scientific explanation distinctively mathematical? *British Journal for the Philosophy of Science, 64*, 485–511.

Lange, M. (2013b). Really statistical explanations and genetic drift. *Philosophy of Science, 80*, 169–188. https://doi.org/10.1086/670323.

Lipton, P. (2009). Causation and explanation. In *The oxford handbook of causation*.

Lyon, A. (2014). Why are normal distributions normal? *British Journal for the Philosophy of Science, 65*, 621–649.

McElreath, R. (2015). *Statistical rethinking: A bayesian course with Examples in R and stan*. Boca Raton: Chapman and Hall/CRC.

Morrison, M. (2014). In *Reconstructing Reality: Models, mathematics, and simulations* (1 edition). Oxford University Press.

Pincock, C. (2007). A role for mathematics in the physical sciences. *Noûs, 41*, 253–275.

Quetelet, A. (1835). *Sur l'homme et le développement de ses facultés ou essai de physique sociale*. Bachelier.

Radick, G. (2011). Physics in the Galtonian sciences of heredity. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 42*, 129–138.

Sober, E. (1980). Evolution, population thinking, and essentialism. *Philosophy of Science, 47*, 350–383.

Stigler, S. M. (1990). Reprint edition. In *The History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge, Mass: Belknap Press.

Stigler, S. M. (2010). Darwin, Galton and the statistical enlightenment. *Journal of the Royal Statistical Society Series A: Statistics in Society, 173*, 469–482.

Strevens, M. (2011). *Depth: An account of scientific explanation*. Cambridge, MA: Harvard University Press.

Strevens, M. (2012). Précis of depth. *Philosophy and Phenomenological Research, 84*, 447–505.

Strevens, M. (2016). How idealizations provide understanding. *Explaining Understanding New Perspectives from Epistemology and Philosophy of Science*, 37–49.

Walsh, D. M., Lewens, T., & Ariew, A. (2002). The trials of life: Natural selection and random drift. *Philosophy of Science, 69*, 429–446.