Introduction

Research on the interaction of nature and nurture now spans nearly a century. (Throughout this article, ‘the interaction of nature and nurture’ or just ‘interaction’ is used to refer to the population-level phenomenon wherein different hereditary/genotypic/genetic groups respond differently to various environmental exposures. I will not be discussing individual-level processes, such as various gene expression mechanisms, where one might say that nature and nurture are interacting.) It has evolved from investigations of strains of crops, to whole genotypes of livestock, to specific genes of humans. It has been performed by a variety of scientists, from epidemiologists and evolutionary geneticists, to developmental psychologists and neuroscientists. It has generated results in the context of eugenic debates about sterilization, social debates about race and clinical debates about treating debilitating disease.

What is remarkable, then, is the fact that the debates about the interaction of nature and nurture have remained strikingly uniform during that century. Even as the scientists, the social context and the technology have changed, the same arguments for and against interaction have turned up time and again. In Beyond Versus: The Struggle to Understand the Interaction of Nature and Nurture I aim to explain this phenomenon—to explain why the struggle to understand the interaction of nature and nurture took the form that it did over the past century and how it continues to shape contemporary research on interaction. The book tells the history of debates regarding research on interaction, draws on tools from the philosophy of science to weigh in on those debates and also looks ahead to ethical discussions regarding genetic testing for genes implicated in gene-environment interactions. For the purposes of this target article, I will focus on those chapters which I believe to be of most interest to the International Journal of Epidemiology readers—the chapters from part 1 which trace the history of the debates, from arguments regarding eugenics in the 1930s to disputes between contemporary researchers regarding the causes of depression.

Interaction in the eugenics controversy

RA Fisher was the first scientist to struggle with the interaction of nature and nurture. Fisher, a giant of 20th century science, formulated in the early 20th century (alongside Sewall Wright and JBS Haldane) the field of population genetics; he also created many of the statistical methods that epidemiologists continue to use today, such as the analysis of variance, the design of experiments and statistical significance testing. Fisher was also an ardent eugenicist. His statistical innovations were developed in part to make eugenic assessments of the relative importance of nature and nurture when it came to evaluating traits like feeblemindedness, and he advocated using that scientific information to intervene in the world with methods like sterilization. Fisher introduced the analysis of variance to provide a tool for partitioning the total variation for some trait of interest into the portion attributable to hereditary variation and the portion attributable to environmental variation. He quickly realized, however, that if the hereditary variation was dependent on the environmental variation (and vice versa), then the components could not be disentangled so easily. Fisher tested for this possibility first with different varieties of potato exposed
to different fertilizer treatments, and when he performed the analysis of variance on the variation in weight, he added an extra component—the ‘deviations from summation formula’—to account for the ‘great complication’ that would present itself if such a problem existed. The deviation from summation was the measure of interaction, and in the potato-fertilizer study it was not greater than chance, so he felt confident dismissing the problem there and in future discussions of variation-partitioning, be it concerning the weight of potatoes or the mental health of humans.\(^5\)

Lancelot Hogben saw things differently. Hogben, though not so well known as Fisher, made equally lasting contributions to science. A trained statistician and biologist, Hogben developed the African clawed frog as a model organism for both experimental research and pregnancy testing; he also founded Britain’s Society for Experimental Biology, which still exists today.\(^6\),\(^7\) Hogben was also a fervent anti-eugenicist. He saw in interaction a devastating critique of eugenic conclusions about fitter or less fit classes of people. The ‘interdependence of nature and nurture’, as he called it, was a developmental phenomenon pertaining to the way in which different genetic constitutions related to environmental exposures during the developmental process.\(^8\) Hogben argued that such cases of interdependence were quite common in nature and should be sought out because of the important information that they shed on elucidating the mechanisms of development. Moreover, he pointed out that a gap in, say, mental health between lower and upper British classes did not necessarily mean environmental intervention was futile, as the eugenicists supposed; for, if interaction existed, different environments could generate vastly different results.\(^9\)

Fisher, when challenged by Hogben on the importance of interaction, was unimpressed. A scientist might study ‘non-linear interaction’ because of some ‘academic’ interest, but it was of no ‘practical’ concern. This was because (i) the phenomenon was rare in nature, and (ii) even if it did turn up in a study, it could be eliminated by altering the scale on which the environment was measured (i.e. a transformation of scale). (Fisher’s reply to Hogben came in the form of correspondence between him and Hogben in February 1933. For the full text of their exchange, see pp. 28–33 of Beyond Versus.)\(^1\)

Fisher and Hogben were, in this way, the first scientists to debate interaction. And they brought radically different perspectives to the exchange. Fisher understood interaction to be a statistical phenomenon—a product of his analysis of variance; moreover, he conceptualized it as an absence—a deviation from summation, a lack of linearity in the formula, a lack of additivity between the main effects. When it came to the question of investigating interaction, Fisher believed interaction had to be considered because it posed a great complication to partitioning variation; however, there was a solution if it turned up—a transformation of scale which made the nuisance go away. Even more fortunately, as far as Fisher was concerned, the transformation of scale was not commonly needed for the simple reason that the phenomenon was rare in nature; it was, he judged, a ‘possible, but unproved’ obstacle to partitioning variation.

Hogben, in contrast, understood interaction to be a developmental phenomenon which he conceptualized as a presence—a presence of a unique source of variation generated by the developmental relationship between nature and nurture. Furthermore, Hogben did not believe interaction was to be ignored or eliminated; rather, the interaction of nature and nurture was to be sought out because of the information it provided about the causal mechanisms of the developmental process. And, finally, because interaction was a developmental phenomenon, it was to be commonly expected in nature.

**Interaction in the IQ controversy**

The next major nature/nurture controversy in which interaction figured began in 1969. By that time, Fisher was dead and Hogben was retired, but their distinct perspectives on interaction arose again. In 1969, educational psychologist Arthur Jensen published ‘How much can we boost IQ and scholastic achievement?’\(^10\) At that time, the most common explanation of the gap in IQ scores between African Americans and Caucasian Americans (about 15 points) was the environmental difference between the two populations; one of them had been exposed to systematic discrimination, segregation and disenfranchisement for centuries in the USA, whereas the other had not. Jensen, drawing on quantitative genetic studies that put the heritability of IQ at about 80%, surmised that environmental differences could not completely explain away the IQ gap. Genetic differences between the races, he countered, have to be considered. This was the birth of the IQ controversy.

Jensen’s genetic hypothesis was attacked from a number of different angles. The data upon which Jensen relied to put the heritability of IQ at 80% were charged with fabrication; IQ tests, it was argued, did not measure intelligence but instead only measured how well people performed on IQ tests; environmental interventions designed to eliminate the IQ gap (like the Head Start programme) had not been given enough time to succeed.\(^11\) Two scientists at Harvard, in contrast, sought to undermine the very methodological foundation of Jensen’s conclusion. Richard Lewontin (an evolutionary geneticist) and David Layzer (a cosmologist) both drew on the interaction of nature and nurture (or genotype-environment interaction when they were writing)
to formulate their assaults. Even if the heritability of IQ was quite high, Lewontin granted, that did not mean environmental intervention was futile; for, as Hogben pointed out decades earlier, a new environment could erase an existing gap if interaction for the trait existed. Analyses of variance that ignored interaction effects were not analyses of causal mechanisms, and that was their fundamental flaw. Such analyses of variance assumed additivity when in fact complex biological traits arose from a complicated developmental process where genetic and environmental factors were intermingled. In short, no inferences about the immutability of the gap in IQ scores between African Americans and Caucasian Americans could be derived from the existence of a heritability score alone (no matter how high).

Jensen, however, was undeterred by Lewontin and Layzer’s emphases on interaction. Like Fisher before him, Jensen countered that interaction was rare in nature. If it existed for a trait like IQ, then an analysis of variance would reveal it; but no such evidence presented itself. Moreover, he complained, appeals to interaction were plagued with confusion; genotype-environment interaction was a statistical concept about individual differences and had nothing to do with individual development. His critics, Jensen replied, put too much emphasis on what might happen in possible environments, when his attention was on what was actually happening in actual environments.

Jensen on the other hand, and Lewontin and Layzer on the one hand, and Lewontin and Layzer on the other, thus found themselves offering up arguments remarkably similar to those that came from Fisher and Hogben decades earlier. Jensen conceptualized interaction just as Fisher had; interaction was a purely statistical concept—an absence of additivity in Fisher’s analysis of variance. Likewise, according to Jensen, the interaction of nature and nurture was not a thing to be studied for its own sake; rather, it was an obstruction to assessing a question like ‘how much can we boost IQ and scholastic achievement?’ which was answered with a partitioning of the genetic and environmental contributions to variation in IQ and scholastic achievement. And, as far as Jensen was concerned, he did not have to worry about the problem because the burden of proof when it came to interaction was on the proponents, and they had failed to provide evidence of the phenomenon in IQ.

Lewontin and Layzer, in contrast, followed Hogben’s lead. Interaction, for them, was ultimately a developmental phenomenon—a product of the developmental relationship between nature and nurture. The genotype, the environment and the phenotype could not be treated as isolated units, the first adding to the second to create the third; rather, the first and the second interacted continuously throughout development, and the third was the manifestation of this interactive, developmental process. Differences in the phenotype, then, would result from differences in this interactive, developmental process. And rather than treating interaction as a nuisance (or as ‘an uninvited party guest’, as Layzer put it), the interaction of nature and nurture should be sought out by constructing norm of reaction charts depicting how different genotypes respond to different environments, since such information shed light on the causal mechanical relationship between the variables. Finally, when it came to the empirical evidence for interaction, Lewontin pointed his readers to known cases of the phenomenon in plants and non-human animals; gathering such results would be much more difficult in humans, he admitted, but since interaction was developmental in nature and known to exist elsewhere it would be the rule and not the exception, and so the burden of proof was placed on those scientists who assumed that additivity between genotype and environment was an accurate reflection of biological reality.

**Interaction in the serotonin transporter gene controversy**

In the early 1990s, scientists began using a combination of quantitative linkage/association studies along with molecular technologies to study the relationship between differences in individual genes and differences in phenotype. The initial batch of findings generated a great deal of excitement, as complex human traits like depression were linked up with specific stretches of DNA like the promoter region of the serotonin transporter gene (5-HTTLPR). That initial excitement, however, quickly gave way to disappointment when follow-up studies failed to replicate the earlier gene-trait findings. Two strategies emerged soon after in response to this dilemma, both of which built off the idea that the fault in the earlier approach was in relying on single genes to explain complex traits. One strategy was the genome-wide association study (GWAS), initially conceived by Neil Risch and Kathleen Merikangas, which sought to identify the sources of genetic variation across the entire genome. The other strategy was the gene-environment interaction study, deployed by Terrie Moffitt and Avshalom Caspi; they combined earlier candidates from the genetic side, like 5-HTTLPR, with an environmental factor like exposure to stressful life events, to capture more variation in depression than either of those factors alone. The gene-environment interaction strategy was the first to generate results; when Caspi and Moffitt’s studies were published, the results garnered a great deal of media attention, and some commentators said their strategy paved the way for a new paradigm in human genetics. The GWAS strategy, in contrast, produced results just a couple of years
later and quickly became associated with the problem of ‘missing heritability’ because the genetic variation found across the genome did not add up to the heritability estimates already established for the traits investigated.20,21

This particular series of events produced an interesting twist in the history of the interaction debates. In previous iterations, studies that highlighted the genetic contributions to human traits came first, and then interaction was introduced by the critics of those studies as a challenge. But this time the tables had turned. Moffitt and Caspi’s studies were not published as challenges to earlier work; rather, they were introducing their own empirical results of interaction. And the critics of interaction, in turn, were not defending themselves; they were now on the offensive. Still, despite the new orientation, the arguments against interaction echoed the earlier replies of Fisher and Jensen. Lindon Eaves, one of the first to question the growing excitement about Moffitt and Caspi’s gene-environment interaction studies, pointed out that variation due to interaction was typically quite small and could be made to appear or disappear with a transformation of scale—the very reply Fisher leveled against Hogben.22 And Stanley Zammit, Michael Owen and Glyn Lewis, in an article devoted to misconceptions about gene-environment interaction, warned that interaction was a statistical effect that had nothing to do with the developmental process—the very point Jensen deployed against Lewontin and Layzer.23

Moffitt, Caspi and allies such as Michael Rutter saw things differently. Research on interaction was not a statistical nuisance to be made to disappear with a transformation of scale; rather, it was designed to inform neurobiological efforts at unravelling the mechanisms of a trait like vulnerability to stress.24,25,26 Moreover, those neurobiological efforts were already well under way. Neuroscientists had designed and performed a number of experiments to test how humans with different variants of the serotonin transporter gene responded to stress exposure, and others designed stress-inducing experiments on non-human animals with an analogue of the serotonin transporter gene.27,28,29

The latest debate over interaction made headlines when a series of three meta-analyses were published between 2009 and 2011. All three meta-analyses were prima facie investigations of the same phenomenon—the relationship between the serotonin transporter gene, exposure to stressful life events, and the subsequent development of depression, as first reported in Caspi et al.19 Marcus Munafò, an advocate of the GWAS strategy, led a team from the University of Bristol (joined by Glyn Lewis) that included five studies in their meta-analysis; that meta-analysis returned a negative result.30 Merikangas, again one of the initial conceivers of GWAS, led a team from the National Institute of Mental Health (joined by Neil Risch and Lindon Eaves) that included 14 studies in their meta-analysis; that meta-analysis also returned a negative result.31 Srijan Sen, who had undertaken earlier research on gene-environment interaction, led a team from the University of Michigan that included 54 studies in their meta-analysis; that meta-analysis returned a positive result.32 This was a strange place for the field to find itself in. Since the three meta-analyses were seemingly all investigating the same interaction effect, it was not at all clear how to interpret the conflict. Indeed, the journal that published Sen’s meta-analysis opted to include a commentary with it, and there the authors worried, “The reader is therefore entitled to ask, ‘What should I believe? Which explanation is true?’ Unfortunately, the answer is unclear, and a long time will pass before questions can be resolved because all the studies so far can be interpreted in opposing ways”.33

### Explaining the interaction debates

What explains these debates? Over the years, a number of historians, philosophers, sociologists and scientists who have themselves participated in the arguments have attempted to answer this question. And the typical answer has consisted in appeals to various socio-political inclinations or lack of expertise on the part of one side or the other. When it came to the debate between Fisher and Hogben, for example, historians pointed to Hogben’s socialism and Fisher’s religiously-inspired brand of eugenics.34,35 Turning to the IQ controversy, Jensen was accused of everything from racism to a basic lack of understanding of statistics, and Jensen’s defenders argued that a whole generation of philosophers has been blinded by political motivations and sloppy thinking.36,37,38 The competing meta-analyses and interpretations of them experienced a similar phenomenon. Sen’s positive meta-analysis was attacked by critics of the gene-environment interaction strategy for being poorly designed because it was too permissive when it came to which replications to include for analysis.39 At the same time, Merikangas and Munafò’s negative meta-analyses were attacked by advocates of the gene-environment interaction strategy for biasing their results by including more negative replications than positive ones.40,41 In sum, each episode has been explained by some form of a divide between one side and the other: a political divide between Fisher and Hogben, a racist divide between Jensen and his critics and an intellectual divide between those who know how to design meta-analyses and those who do not.

The problem with each of these explanations is that they all attend to just one specific episode of the interaction...
debates and get focused on the particular social context, the particular technology and the particular players in that episode. But, as mentioned at the introduction and conveyed in the previous sections, the arguments for and against interaction have remained constant despite the changes in social context, technology and cast of participants. If Jensen’s arguments against interaction were just a product of racism, then why do those same arguments turn up again in the 21st century when race has nothing to do with the serotonin transporter gene controversy? And if Hogben’s arguments in favour of interaction were derived from socialism, then why do those same arguments turn up again in the 21st century when political concerns about the economically disadvantaged play no role in the current debate?

What we need to explain the interaction debates, then, is an account of some divide that spans the past century, a divide that is not reliant on the particulars of one episode but instead captures them all. An explanatory divide provides that solution. Throughout the history of genetics, an explanatory divide has existed between those who employ a variation-partitioning approach and those who employ a mechanism-elucidation approach to the science. Both sides share a common interest in studying the relationship between nature and nurture, but that commonality belies the fact that they go about their studies in two very different ways. They identify different things that need explaining; they ask different questions about that thing that needs explaining; they point to different things that do the explaining; and they utilize different methodologies to provide those explanations. For variation-partitioners, the thing that needs explaining is variation in a population. They ask ‘how much?’ questions about that variation. They seek to identify and measure the causes of variation responsible for the variation. And they employ methods like the analysis of variance to supply those answers. For mechanism-elucidators, in contrast, the thing that needs explaining is the developmental process. They ask ‘how?’ questions about that process. They seek to elucidate the causal mechanisms responsible for the outcome of that process. And they employ experimental methods that intervene in the mechanisms to generate that elucidation.

So there is a divide between variation-partitioners and mechanism-elucidators, but it cannot be explained completely by appeals to politics, biases or intellectual aptitudes. There is also a divide concerning how explanation works in science. And depending upon which side of the explanatory divide a scientist falls, the interaction of nature and nurture figures quite differently. For the variation-partitioning approach, where the whole goal is to disentangle and then measure the various causes of variation responsible for variation in a population, the interaction of nature and nurture presents a genuine obstacle to that endeavour. For the mechanism-elucidation approach, where the whole goal is to probe and understand the causal mechanisms responsible for the developmental process, the interaction of nature and nurture presents a genuine insight into that endeavour. We can see the explanatory divide in each episode of the interaction debates.

Fisher was the architect of the variation-partitioning approach. He created the statistical tools that allow a scientist to evaluate variation in a population and measure the causes of that variation. The absence of additivity in those causes of variation, he quickly realized, was a serious problem for his methods, and so his struggle with the interaction of nature and nurture was a struggle with minimizing it. Hogben, an advocate of experimental biology, favoured a mechanism-elucidation approach to studying nature and nurture. The presence of interdependence between nature and nurture, he pointed out, was an important piece of information if a scientist wanted to understand how those variables interacted during an organism’s development, and so his struggle with the interaction of nature and nurture was a struggle with making the most of it. Now, this isn’t to say that politics played no role in the exchange between Fisher and Hogben during the eugenics controversy; they were, after all, people—people with political persuasions and ethical convictions and attitudes about religion. But it is to say that we need not appeal to those facets of their individual personas to make sense of their disagreement about interaction.

Jensen operated in the variation-partitioning approach as well. He employed the standard variation-partitioning measure of his time—heritability, and he asked a standard variation-partitioning question—how much can we boost IQ and scholastic achievement? Focused on that question and attempting to address it with a heritability score, the interaction of nature and nurture presented a serious problem for answering it. Lewontin and Layzer endorsed a mechanism-elucidation approach to studying nature and nurture. If intervention and prediction were the goals, they counselled, then we had to know how a system works, and that meant understanding the causal mechanisms of cognitive development that gave rise to something like an IQ score. Attention to the interaction presented in a norm of reaction graph, though no blueprint of the causal machinery, at least provided some information as to how two of the variables were causally related. Again, this is not to say that socio-political concerns had no influence in the IQ controversy; the debate over race, race differences and the causes of those differences was highly charged, and the disputants certainly had their eyes on the ethical, social and political implications of the conclusions. But it is to say
that we need not appeal to those elements of that particular controversy to make sense of the disagreement over interaction.

That brings us to the present-day serotonin transporter gene controversy. The GWAS strategy is the modern-day exemplar of the variation-partitioning approach. Administrators of a GWAS, like designers of heritability studies before, focus on measuring the genetic causes of variation, determining how much of that genetic variation accounts for the total phenotypic variation under investigation. The advocates of the gene-environment interaction approach, in contrast, have put much more emphasis on how their research informs the search for the mechanisms of trait development. This contrast was on full display when Moffitt and Caspi replied to the two negative meta-analyses by drawing attention to the experimental human and non-human studies that supported the idea that there was some neurobiological mechanism implicating the serotonin transporter gene in a sensitivity to stress exposure.\(^40,42\) The designers of the negative meta-analyses, in reply, simply said that the experimental research was irrelevant to the matter at hand.\(^43,44\) The contrast between approaches is also manifest in the disagreement about which of the meta-analyses were designed correctly or incorrectly—the negative meta-analyses which included relatively few studies, or the positive meta-analysis which included many more. Designers of the negative meta-analyses were very strict in terms of which studies actually counted as a replication, requiring the studies to, in particular, measure the environmental variable just as Moffitt and Caspi had (with exposure to multiple stressful life events grouped together and then sequentially added). But this inclusion criterion then excluded many studies that examined how individuals with different variants of the serotonin transporter gene responded to specific stressors, like hip fractures, heart attacks and strokes. Designers of the positive meta-analysis included these. So the question becomes: which design is correct? Whereas participants in this debate have weighed in on this question by making the case for one side or the other, I argue instead that the explanatory divide on which you fall provides different warrants for favouring the results of either of the meta-analysis designs.

**Conclusion: the advantage of an explanatory divide**

The nature/nurture debate generally and the interaction debates specifically have been contentious. Accusations of bias, racism and just plain ignorance have pervaded the territory since its beginning. My appeal to an explanatory divide, I claim, is both historically accurate and, more importantly, scientifically and philosophically productive. It is historically accurate because it better accounts for the fact that the same arguments for and against interaction have been deployed for nearly a century now, defying explanations that appeal to the idiosyncrasies of individual episodes. It is scientifically and philosophically productive because, unlike charges of bias, racism and ignorance, an explanatory divide calls out for an explication of the relationship between the variation-partitioning and the mechanism-elucidation approaches to studying nature and nurture. And that is something from which all sides can benefit.

**Conflict of interest:** None declared.

**References**