9-3-2015

What difference does it make? An essay review of Beyond Versus: The struggle to understand the interaction of nature and nurture; James Tabery; MIT Press: Cambridge, MA, 2014

Peter J. Taylor
University of Massachusetts Boston, peter.taylor@umb.edu

WORKING PAPER on Science in a Changing World

Paper # 1-2015

What difference does it make?

An essay review of Beyond Versus: The struggle to understand the interaction of nature and nurture; James Tabery; MIT Press: Cambridge, MA, 2014

PETER J. TAYLOR

GRADUATE PROGRAM IN CRITICAL AND CREATIVE THINKING
University of Massachusetts, Boston, MA 02125, USA
http://scholarworks.umb.edu/cct_sicw/1

What difference does it make?
An essay review of Beyond Versus: The struggle to understand the interaction of nature and nurture; James Tabery; MIT Press: Cambridge, MA, 2014[1]

Peter J. Taylor
Science in a Changing World graduate track
University of Massachusetts, Boston, MA 02125, USA
peter.taylor@umb.edu

Abstract

Beyond versus makes its contribution to the thriving industry of books that clarify or recast nature-nurture issues through seven conceptual moves. The first is to posit a divide between sociological and philosophical inquiry. As Tabery depicts them, commentators on the science invoked in nature-nurture debates often focus on the racist or other political views of disputants or on their flawed understanding of scientific concepts. Tabery, in contrast, as a philosopher of science, explains past and present disagreements as stemming from "a disagreement concerning how explanation works in science." (The other moves include explanatory and terminological divides, connecting associations to mechanisms, rank-change versus divergence-only interaction, a single category for nature-nurture.) This review essay, while operating for the most part on the philosophical side of the divide, does promote more careful understanding of the science of data analysis. This leads me to present alternatives to each of Tabery's moves, including, eventually, the sociological-philosophical divide.

---

[1] A shorter version of this paper has been published as Taylor, Peter J. (2015) "Distinctions that make a difference? (An essay review of Beyond versus: The struggle to understand the interaction of nature and nurture by J. Tabery)," Studies in History and Philosophy of Biological and Biomedical Sciences, Part C, 51: 70-76.

Genetic variants, definitions of interaction, philosophical concepts, analysis of observations versus experiments, a focus on agricultural versus human subjects, a book review…—all these things may make a difference. Whether they do depends on other things, which may or may not be controllable. I tease out the potential significance of these two opening sentences in this essay review of *Beyond Versus*. Along the way, I identify and assess seven key conceptual moves Tabery makes in his 2014 contribution to thriving industry of books that clarify or recast nature-nurture issues (reviewed in Appendix 1).

On Move 1, Dividing sociological and philosophical inquiry

As Tabery depicts them, commentators on the science invoked in nature-nurture debates often focus on the racist or other political views of disputants or on their flawed understanding of scientific concepts. Tabery, in contrast, as a philosopher of science, explains past and present disagreements as stemming from "a disagreement concerning how explanation works in science" (p. 5; from hereon page numbers on their own refer to *Beyond Versus*). This essay operates for the most part on the philosophical side of the divide in order to provide meaningful commentary on Tabery's distinctions and concepts. I do not make sociological or political interpretations of flawed understandings, but I do promote more careful understanding of the science of data analysis. This leads me to present alternatives to each of Tabery's moves, including, eventually, the sociological-philosophical divide.

On Move 2, Explanatory and terminological divides

Tabery proposes that the history of nature-nurture debates is not about scientists contesting whether a given trait is determined by heredity or by the environment, but rather about the significance they to the *interaction* of heredity and environment. The scientists in the three episodes he examines in Part I talk past each other because one side approaches interaction by partitioning variation, while the other seeks to elucidate mechanism (Table 1, drawn from p. 124). The second of his key conceptual moves is

to identify this explanatory divide; understanding it enables us, Tabery argues, to move *beyond versus* as well as gain insight into some bioethical issues arising in this era of gene-based diagnosis. The alternative I suggest is to view the relevant sciences as united under the variation-partitioning approach (Table 2) while divided by the meanings given to interaction, which are as different as chalk, cheese-sticks, and lipstick. The basis for this alternative view, which draws on my formative research experience in the 1970s analyzing data from large plant breeding trials (Taylor 2014a), needs to be laid out before Tabery's three episodes can be examined.

Table 1. The Components of the Explanatory Divide (Tabery)

	Variation-partitioning approach	**Mechanism-elucidation approach**
Thing to be explained	Variation in a population	Developmental process
Causal question	How much?	How?
Thing that does the explaining	Cause of variation	Causal mechanism
Methodology	Statistical	Interventionist

Table 2. The Common Components of Analysis (this essay)

	Variation-partitioning approach
Thing to be analyzed	Variation of an observed trait in a population in a range of situations
"Causal" question	How much difference in an observed trait is associated with differences in other things?
Thing that accounts for the difference	Thing that is significantly associated with the variation in the trait
Methodology	Statistical

The science reviewed in *Beyond Versus* centers on the statistical analysis of data sets in which some given trait varies across a population of individuals (which may be people, plants, fruit flies, and so on) of various degrees of relatedness, raised in

various situations (which may be families, geographic locations, or specific conditions, e.g., plants grown with 50 kg/ha of nitrogen fertilizer). The analyses are primarily of *observational* data, which is derived from individuals that can be subdivided into relevant categories (e.g., people raised in low socioeconomic status), and only in a few instances of *experimental* data, which involves assigning individuals randomly to be subject to specific conditions. Sometimes the categories into which individuals are subdivided are defined by genealogical relatedness without knowledge of measurable genetic factors that underlie the relatedness or by location without knowledge of the underlying environmental factors in each location. At other times the categories are defined by measured factors, such as presence or absence of a specific genetic mutation, socioeconomic status, amount of fertilizer applied, and so on. (*Factor* is used in this essay in a non-technical sense, referring simply to some thing whose presence or absence can be observed or whose level can be measured.)

Statistical analysis of data connects the observations of a given trait to a model that is static (in contrast to dynamic models such Newtonian equations governing bodies in orbit). The models relevant to *Beyond Versus* can be divided into three types (Figure 1; expressed in equations in Appendix 2):

A) a summation of variables (technical name *effect*) derived from the observations (e.g., the value for the trait in a certain plant variety averaged over all the locations in which it is grown);

B) a summation of measured genetic and environmental factors, each weighted by a coefficient; and

C) a summation that combines features of A and B and in which the environmental factor is experimentally manipulated.

The summation for each kind of model also includes a non-systematic residual contribution. By adjusting the details of the model (e.g., for type B, the values of the coefficients), the discrepancy between the observed values for the trait and the prediction of the trait's value based on the model can be minimized.

Figure 1. Three kinds of model for statistical analysis of observations on traits

		Location/family (A), Measured environmental factor (B & C)				
		1	**2**	**3**	**4**	**etc**
Variety/twin pair (A & C), Measured genetic factor (B)	**1**					
	2		*trait values*			
	3					
	etc					

Trait value for an individual in a given cell*	= overall average in data set for the trait + contributions for the Row + Column + Row-column-combination + Residual
Row contribution	for A & C = average over all the locations in which variety is raised** for B = measured genetic factor weighted by a coefficient
Column contribution	for A = average over all the varieties raised in that location for B & C = measured environmental factor weighted by a coefficient
Row-column-combination contribution	= average over individuals in that cell – contributions for Row & Column
Residual	= difference between trait value for an individual and the sum of the above contributions

* Some or many cells may be empty. For example, in studies of human twins, members of each pair are raised in at most two families. ** In practice, genealogical relatedness of individuals is also taken into account in estimating these contributions.

The variation in the trait can be subdivided or *partitioned* into components associated with the different variables or measured factors in the models as well as the variation of the residual discrepancies. Partitioning of variation, whichever type of model is used, always entails a "how much" question in that statistical analysis assesses which components of the trait variation are significantly greater than the residual. When a component is not significant, the model is reformulated without the corresponding terms. (Elaborations on these basic types and the technicalities of how statistical significance is assessed do not affect the conceptual points made in this essay.)

In particular, statistical analysis of the partitioned variation assesses whether there is significant variation associated with *interaction*, the statistical term for the row-

column-combination contributions in Figure 1. The alternative view makes the following features of interaction clear:

- Interaction in statistical analysis is not dynamic in the sense, say, of two soccer players contesting control of the ball.
- Interaction is not synonymous with interdependence in a colloquial sense given that, even if interaction were zero, each kind of contribution is *conditional* on the full set of individuals and situations where they are observed. (For example, in model A, the contribution of a variety/twin pair is not a property of the variety/twin pair, but is the value for the trait averaged over *all the particular locations/families in which it is raised*. Change the set of varieties and locations in which the trait is observed, the size or even significance of the associations may change, as Turkheimer at al. 2003 illustrates. Similarly, in type B analyses, expand or contract the range of factors in which the trait is observed, the size or even significance of the associations may change.)
- All three kinds of model are simple sums whether or not the interaction contributions are significant (which is good reason to avoid the term *non-additivity* that Tabery, p. 22, following some researchers, uses to describe the presence of a significant interaction).
- No conceptual or empirical connection exists between the terms in the different models, including the corresponding kinds of interaction, because the different kinds of analyses involve different *things*—variables derived from the observations versus measured factors.

Let me address two objections to this last point. The first possible objection: a gradient of measurable, albeit yet-to-be-identified factors might run through the variables derived from observations in type A and C analyses. This, however, need not be the case, which is obvious when we think about, say, human height. Pathways of development involving diverse combinations of genetic and environmental factors make intuitive sense when we note the different timing of growth and the make-up of the final height (e.g., long trunk, short legs versus short trunk, long legs) (Taylor 2014a, 19, 28ff). The lack of conceptual or empirical connection between measured *factors* and variables derived from the observations of *traits* is especially relevant in discussions of *heritability*.

This is the technical term for the variation among the row contributions in type A analysis as a fraction of the total variation and has *nothing* to do with the colloquial view that a trait is heritable when it involves transmission of a gene or genetic factor from parent to offspring. (This regrettable ambiguity is amplified when researchers who are proficient in type A analysis refer to heritability as the "contribution of genetic differences to observed differences among individuals." The quote is from Plomin et al. [1997, 83], but the interpretation is widespread and is repeated by Tabery, p. 47, 92; see Taylor 2014a, 24ff. The points noted briefly in Appendix 3 accentuate why the interpretation is misleading. Similarly, interpretation of other fractions of variation in terms of differences in yet-to-be identified environmental factors is not warranted.)

The second possible objection: measured environmental factors are involved in both type B and type C analyses. However, the measured factors in type B analyses need not be *modifiable* (e.g., chromosomal sex is a commonly measured but non-modifiable genetic factor). Moreover, if the factors were modifiable, *it does not follow that modifying them would generate the differences observed in the original data set*. In other words, it does not follow that the difference that "makes" a difference as exposed by statistical analysis of observational data is a factor we can modify to make the same difference again. For example, lower income level is a significant factor associated with smoking rates, but there is no reason to expect that disbursing $10,000 to poor smokers would lead many of them to quit. After all, the dynamics through which a person develops a low income and the dynamics through which a person becomes a smoker are separately and jointly far more complex than any static statistical model can capture. For the variables in type A analysis, as well as for all terms in type C analysis other than the modifiable environmental factors, the point on conditionality above means that it is not possible to undertake an *intervention* to change "the thing associated with the variation." In light of this and because measured factors in type B analyses are not necessarily modifiable, Table 2 places "causal" in scare quotes and substitutes the conventionally ambiguous statistical term "accounts for" for Table 1's "does the explaining."

The points made in the preceding paragraph become salient when we return later to discussion of mechanism. First, let us revisit the historical and current debates

examined by Tabery in light of the distinction between the three types of data analysis and the corresponding forms of interaction, where the analysis in each case involves partitioning of variation of an observed trait in a population.

Reinterpreting three debates, historical and current

Opponents in nature-nurture debates about humans are often debating whether or not society should do more to enhance the range of situations in which people develop. In the Jensen-Lewontin debate discussed by Tabery (p. 46ff), Jensen observed that compensatory preschool education programs that began in the late 1960s, such as Headstart, had only a transitory effect on boosting IQ test scores. He saw the high heritability of IQ test scores, in combination with the lack of success in reducing the gap between the average scores for black versus white Americans, as conferring plausibility on the hypothesis that the gap is associated with differences between the races in some yet-to-be-determined genetic factors (Jensen 1969, 1970). In contrast, Lewontin, as Tabery (p. 50) notes, asserted that we can "boost IQ and scholastic achievement [by] as much or as little as our social values may eventually demand" (Lewontin 1970, 25). Lewontin was critical of heritability estimation and, even more so, its interpretation, but one issue he did *not* take up was Jensen's (1969, 39) assertion that the contribution of interaction to variation in "intelligence" is small in relation to other contributions. Lewontin and Jensen were both operating on the variation-partitioning side of Tabery's divide (Table 1), on the terrain moreover of type A analysis, not of measurable genetic and environmental factors. One caveat: Later, the arguments of Lewontin (1974, 1982) made use of examples like those offered by both Fisher and Hogben in a debate forty years earlier.

In the Fisher-Hogben debate of the 1930s (p. 15ff), both researchers examined cases in which different varieties were raised at different levels of an experimentally manipulated environmental factor (type C analysis). Fisher found, for example, in potato varieties subject to various levels of manuring, that the additional contribution from variety-environmental factor combinations was not significant. Hogben, on the other hand, found that the difference in numbers of eye facets in fruit fly strains was sensitive to the temperature at which flies were raised. As Tabery (p. 32) notes, he

drew the lesson that social inequalities could be diminished or exacerbated by the medical, educational, and nutritional environments in which people were raised (echoed 40 years later by Lewontin). In opposing the eugenic sentiments of Fisher, Hogben depicted the methods of partitioning variation that Fisher had been pioneering as the basis for incorrect extrapolation of what has been observed to what was possible. A problem for his critique is that the same variation-partitioning methods could be used for analysis of the data in both cases. Moreover, because the environmental factors could be experimentally manipulated, both cases could go beyond establishing associations to illuminating mechanisms, albeit with the role of the genetic factors underlying the different varieties remaining unknown. If we assume that the researchers had varied the environmental factors across the feasible range, Fisher's trials were showing that the best yielding varieties on average could also get the best out of each level of the environmental factor (manuring). There was no need for recommendations to farmers about what varieties to grow or to plant breeders about which varieties to breed from to be tailored to the specific level of manuring. Hogben's experiments, however, were showing that varieties could be influenced by an environmental factor (temperature) similarly at some levels and divergently at others. In short, establishing what associations (especially interactions) and mechanisms applied in general was not a matter of contrasting concepts or methods—with Fisher and Hogben being on opposing sides of Tabery's explanatory divide—but an empirical matter that depended on the set of varieties, the environmental factors manipulated when raising those varieties, and the trait. Moreover, whatever Fisher and Hogben believed about the possibilities and pathways for improvement in human populations, no such experimental manipulations existed for human traits.

What Fisher and Hogben lacked is now, in Tabery's view, provided by research such as that of Caspi, Moffitt and colleagues on associations of human psychological traits with combinations of measured genetic and environmental factors (i.e., type B analyses). Caspi et al. (2002), for example, reports on antisocial behavior in adults in relation to the activity of monoamine oxidase type A (MAOA) and childhood maltreatment; MAOA deficiency is a strong predictor of antisocial behavior only when the child has also been maltreated (Figure 2). In other words, there is interaction

between the measured genetic and measured environmental factors. The scientific debate about this kind of interaction, discussed by Tabery (p. 87ff), has revolved around meta-analyses assessing the generality of findings of significant interaction associations: Caspi et al. (2003) is confirmed by one study, but not by two others. Tabery notes that authors of the latter meta-analyses are advocates of using Genome-Wide Association (GWA) studies to detect associations between traits and multiple genetic variants. Tabery positions GWA studies on the variation-partitioning side of his explanatory divide, whereas he places research to detect association of traits with combinations of measured genetic and environmental factors on the mechanism-elucidation side. Under my contrasting view, both sides of the meta-analysis debate operate on the variation-partitioning side, this time on the terrain of type B analysis.

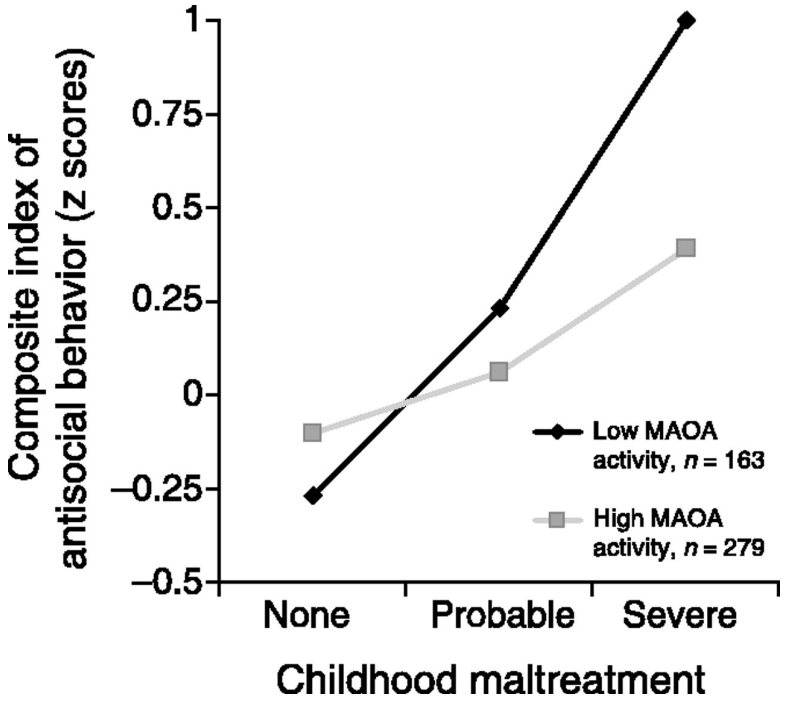

Figure 2. Average adult composite antisocial behavior score in relation to levels of MonoAmineOxidaseA and level of childhood maltreatment for a sample from Dunedin, New Zealand (from Caspi et al. 2002, 852, reproduced with permission).

In summary, none of the three scientific episodes involve disputants fundamentally divided by how to detect interaction even if they differ in the amount of

interaction found in their data sets. However, from one episode to the next, we see a different kind of data analysis and meaning of the term interaction (Table 3).

Table 3. The significance of "interaction"* contributions in three scientific episodes reviewed by Tabery

	Partitioning variation		
	Observational data		**Observational-experimental data****
Episode	**A. summation of variables derived from the observations**	**B. summation of measured factors**	**C. Hybrid of A and B**
Fisher vs. Hogben			Contingent empirical matter for non-humans
Jensen vs. Lewontin	(At the time) Jensen's claim about lack of significance not a point of contest.		(Later) Lewontin uses examples like Fisher's and Hogben's to make his arguments.
Caspi et al. vs. GWA studies		Disputed empirical matter	

* See text for elaboration of the points about the significance of interaction in the episodes. The scare quotes are placed around "interaction" to emphasize that the term has different meanings in each of the three kinds of analysis. ** Observed varieties and experimentally manipulated environmental factors

Moves 3-5, Connecting associations to mechanisms

> [A] scientist explains a phenomenon by identifying and manipulating the variables in the mechanisms responsible for that phenomenon, thereby determining how those variables are situated in and make a difference in the mechanism (p. 109).

The concept of mechanism, much discussed by philosophers of biology during the last decade or more, is the focus of Part II of *Beyond Versus*. Suppose that statistical analysis of observations has identified variables or factors as associated with a given trait. In light of the quote above, it is clear that, in order to connect these variables or factors to a mechanism, they have to be manipulable. In the case of type C analysis,

the environmental factors can by definition be modified, but this is not obviously the case for the other variables and contributions in the three models used to assess associations.

Tabery clearly imagines that measured genetic and environmental factors can be manipulated when, in the third of his key conceptual moves, he positions the type B analysis of Caspi and colleagues on the mechanism-elucidation side of his explanatory divide and labels the method interventionist. Indeed Caspi et al. (2002, 853) conclude that their results "could inform the development of future pharmacological treatments." The implication, in the context of research on childhood experience in relation to adult behavior, is that, if low MAOA children could be identified, prophylactic drug treatment could reduce their propensity to antisocial behavior as adults. To be more precise—and to highlight the interaction component—such treatment could reduce their vulnerability to childhood maltreatment in the sense of the risk that maltreatment would pave the way to undesired adult outcomes. An easy rejoinder (as Tabery, p. 183-5 notes) would be that, if childhood maltreatment could be prevented, children's low MAOA levels would no longer make them more likely to end up as antisocial adults. Some authors, reviewed by Tabery (p. 173ff), have proposed monitoring and measures to prevent maltreatment for children diagnosed at birth as low MAOA or want to avoid the problem altogether by pre-implantation genetic diagnosis and elimination of low-MAOA embryos.

The preceding actions all depend on linking statistical associations to inquiry into mechanisms by viewing the measurable factors as ones that can be manipulated. As a general perspective, this view is not warranted. As noted earlier, the measured factors in type B analyses need not be modifiable and, if they were, it does not follow that modifying them would generate the difference observed in the original data set. Measured factors that are statistically significant but *not* modifiable serve as an invitation to researchers to probe further and try to expose underlying factors that might be modified. For example, the higher incidence among African-American women of pre-term delivery of their babies has shown to be associated with self-reported experience of racial discrimination even after allowing for other factors associated with that outcome—alcohol and tobacco use, depression, education, and income (Mustillo et al. 2004). Perhaps, self-reported experience of discrimination is itself associated with

further underlying factors, so that, as in the earlier smoking example, if there were policies to *directly* counter experiencing discrimination, they would not have the effect suggested by the static statistical model. Perhaps, however, investigations that take this specific difference-associated-with-a-difference as an entry point for inquiry might eventually contribute to understandings in which subjective experiences get brought into a dynamic picture of the biological and social developmental processes that lead to pre-term delivery—a picture, moreover, that could inform actions to reduce the disparity in pre-term delivery.

These last sentences are intentionally tentative: It is a contingent matter whether the full picture of mechanisms and processes of development can be pieced together so as to inform possible actions. Let us illustrate such contingency with a key case discussed by Tabery. He notes (p. 121ff) that variation in human populations in a gene labeled BDNF (brain-derived neurotrophic factor) has been shown to be associated with variation in hippocampal activity and with a test of spatial memory. As in the MAOA case, the association has led to proposals for action, in this case, BDNF-enhancing diets (evident by searching "BDNF boost spatial memory" on the internet). To make the connection between this association and mechanisms, Tabery (p. 110ff) points not to such actions, but to experiments in which BDNF is manipulated. The experiments, undertaken on mice, have resulted in a multi-level explanation of spatial memory that spans from changes in BDNF through activity of receptors, long-term potentiation of neurons, maps in the hippocampus, to navigation of mazes. Now, the validity of mice as a model for humans is an issue well recognized by researchers. Yet even to speak of "mice" and "humans" is to adopt a framing that discounts the variation among mice and the variation among humans. If, instead, we were to pay attention to the variation, the first step would be to note that highly selected strains of laboratory mice are less variable than undomesticated populations (Rader 2004) and experiments made on such mice involve tightly controlled situations. To what extent, it might be asked, do experimental observations hold for individuals from undomesticated populations raised in varied and far more complex situations? If mechanisms have been exposed using laboratory mice, to what extent do they depend on the controlled value of factors that are not typically enumerated when describing the mechanism?

To ask such questions is not to counsel despair. Experiments on humans are possible, albeit with varying degrees of control over the subjects and conditions. Most notably, randomized control trials (RCTs) look for an association with a single manipulated factor, such as a drug versus a placebo, against a background of all other factors varying randomly in the population. A newer approach, less well-known, is Mendelian randomization (Davey Smith and Ebrahim 2007), which uses natural experiments to look for an association between, for example, C-reactive protein (CRP) levels in the blood and coronary heart disease (CHD) for people who have a rare genetic variant that leads to life-long elevated CRP levels, but otherwise vary randomly on other risk factors for CHD (such as smoking, bodymass index, and blood pressure). (CRP levels are associated with an increased incidence of diabetes, hypertension, and cardiovascular disease [Ridker et al. 2007], but Mendelian randomization cast doubt on any causal connection [C Reactive Protein Coronary Heart Disease Genetics Collaboration 2011].)

In general, experiments on humans involve, however, less control than in RCTs or Mendelian randomization. It should not be surprising that, as noted earlier, even when the measured factors can be modified, this need not replicate variation in the dynamics that generated the original data and thus that association. Returning to the MAOA case, medication throughout childhood could have side effects that might not emerge until later in life; an experiment that involved pre-implantation elimination of low-MAOA embryos would bring at the very least all the long-term risks of being conceived by in-vitro fertilization (e.g., Hargreave et al. 2013). Detecting and preventing childhood maltreatment might require intrusion into many households, surveillance, and intervention by state agencies, diversion of government budgets from other needs, and so on. Such changes in the way society runs might well have consequences for the development of children, even those whose MAOA levels were not low. The conundrum is that we would have to know a lot about the processes of development of children in their psychosocial context in order to interpret any experiment that sought to translate associations based on type B analyses into knowledge about mechanisms, let alone into insight about developmental processes.

The contingency and complexities of relating type B associations to mechanisms speaks to the fourth and fifth conceptual moves in *Beyond Versus*:

- drawing on the concordance of the BDNF-memory association found in humans with the experimental research on BDNF in mice, call such an association a *population mechanism* and portray these as a bridge between the variation-partitioning and mechanism-elucidation approaches;
- adopt Waters's concept of *actual difference maker*, which holds that, although many genes might *possibly* make a difference to the development of a trait, the cause of a difference in the trait can said to be the difference in the gene that is actually associated with that difference.

These moves are hard to reconcile with this essay's distinctions between type A and B associations, between unmodifiable and modifiable measured factors, and between associations and manipulating modifiable factors to generate the difference observed in the original data set. If we take these distinctions into account, studies such as those of Caspi, Moffitt and colleagues detect associations whose relevance to elucidation of mechanisms is contingent. In other words, a statistical difference maker does not necessarily *make* a difference. The connection between an association in a population and mechanisms is susceptible to disconfirmation by experiments and invites scrutiny of the relationship of experimentally altered dynamics to the dynamics that generated the original data analyzed to show the association. The association is also conditional: understanding it and formulating manipulations based on it requires attention to the other measured factors experimentally or statistically held constant. The understanding and manipulations need not extrapolate beyond the original population and situations. In other words, *possible* difference makers are causally relevant.

An aside on Move 1, Dividing sociological and philosophical inquiry

One reading of *Beyond Versus* is that Tabery is impressed by advances in genomic science, which makes him optimistic about elucidating mechanisms by moving downward or inward into the molecular basis of traits. Yet, interesting mechanisms may also be elucidated by moving upward or outward. Consider human metabolic diseases,

taking phenylketonuria (PKU) as an example (Paul and Brosco 2013). The cognitive development of individuals with PKU is extremely impaired by the level of the essential amino acid phenylalanine present in normal diets. The level of phenylalanine can be manipulated to reduce greatly the impairment. Social support practices can be adjusted to enhance compliance with the diet, as can policies regarding insurance coverage of the diet. Changes in policies and practices regarding contraception and abortion can be investigated in relation to the incidence of so-called maternal PKU—children born to women who did not strictly maintain the diet. Relevant mechanisms can also, of course, be investigated by moving downward/inward. Mutations in the PAH gene were long ago shown to underlie PKU; the possibility of gene therapy, involving manipulation of PAH genes in stem cells, is now being considered; researchers are examining the responsiveness of individuals with different PAH mutations to a drug, BH4, that allows for a higher-protein diet. Yet, such research on molecular genetic and pharmacological mechanisms depends on further upward/outward experiments ranging from social support to discourage individuals who take BH4 from going off the special diet altogether to the implementation of government subsidies for the biotechnology industry.

Now, philosophers of biology might not inclined to expand their research on mechanisms in the upward/outward direction, but we can hardly come to conclusions about which direction makes the most difference—leads to the most impact for the science—without sociological inquiry into the funding of research and adoption of its findings.

Move 6, Rank-change versus divergence-only interaction

Reservations about connecting measured factors with mechanisms have to be put aside in order to review Tabery's bioethical discussion, which makes up Part III of *Beyond Versus*. In the sixth of the key conceptual moves that I identify, Tabery distinguishes gene-environment interactions in which the average response of the two genetic variants changes *rank* across the range of environmental factors from those in which the differences in averages simply diverges sufficiently for the contributions of the gene-

environment combinations in type B analysis to be associated with a significant fraction of the trait variation. Tabery is concerned that, in the rank-change cases, any action to improve the outcome at one end of the range will diversely affect the outcome for the other variant at the other end of the range. For example, as indicated in Figure 2, boosting MAOA for low-MAOA children to reduce their vulnerability to severe maltreatment would increase the average anti-social behavior for low-MAOA children who are subject to no maltreatment.

Tabery's discussion of rank-change cases is based on plots, such as Figure 2, of *averages* for the trait for the different combinations of measured genetic and environmental factors. *Variation around the averages is discounted*, as is also the case when considering prophylactic drug treatment and prevention of childhood maltreatment (mentioned earlier). Such *typological* readings of data can be countered by paying attention to variation and noting that, within each combination of factors, people show a range of antisocial behaviors. Among children who experienced probable or severe maltreatment, the ranges overlap, that is, some of the high MAOA individuals ended up with higher antisocial behavior scores than some of the low MAOA individuals. Once the resources are invested to screen children for MAOA levels, a troubling issue of misclassification would arise given that attention would be focused on *all* low MAOA children. Indeed, how could treating children according to their genetic group be avoided if we do not know from a childhood MAOA assessment whether any particular individual is one who would go on, after maltreatment, to become an antisocial adult? (Taylor 2014a, 132ff). If misclassification is seen as a bioethical concern, it is a concern that applies whether or not interactions are rank-changing.

Move 7, A single category for nature-nurture

The last conceptual move I identify in the book is actually one that Tabery makes at the very beginning of Beyond *Versus*: subsume the different ways that researchers and others invoke hereditary versus environmental influences under the one label nature versus nurture; anything that seems to involve interdependency of those influences becomes a matter of interaction (p. 1ff). The discussion in this essay allows us, in

contrast, to distinguish four disjunct areas of nature-nurture science. First, through type A analyses, researchers can try to compare how much variation is associated with differences among means for varieties, locations, variety-location combinations, and residual contributions (i.e., the ambiguously labeled genotypic, environmental, genotype-environment interaction, and error variance). Second, through type B analyses, researchers can try to compare how much variation is associated with differences in measured genetic factors, environmental factors, gene-environment interaction, and a residual component. Third, either through type A or B analyses, researchers can compare the variation within groups (e.g., among Euro-Americans and among African-Americans) to the difference between the averages for the groups. Fourth, through investigations that might extend any of the preceding kinds of analysis of observational data, researchers can piece together a picture of the processes of development of a trait and, on that basis, speak to the fixity versus flexibility of traits. (A fifth kind of nature-nurture science examines the basis for human traits in an evolutionary past, but that lies outside the issues discussed in this essay.)

In the long history of nature-nurture debates, opposing sides often assume, imply, or propose that these different sciences are speaking to the same issues. This sense of equivalence or, at least, mutual relevance is evident most notably in discussions that create or play on ambiguity in the meaning of the technical term heritability as well as in unwarranted interpretation of other fractions of variation in terms of differences in yet-to-be identified environmental factors. The misinterpretations of heritability—or, more generally, of the relative sizes of different components of variation estimated using type A analyses—may seem moot if they are seen merely as *heuristics* to guide researchers when choosing which traits to investigate further to identify the measured genetic or environmental factors. The use of molecular tools to identify genetic variants associated with variation in traits is illustrated by the BDNF and spatial memory case that Tabery discusses. Yet for human medical traits, the most powerful new approach, Genome-Wide Association studies, has only found associations with genetic variants that correspond to a small increase in incidence of the trait (McCarthy et al. 2008). The hope had been to expose variants corresponding to a major increase in incidence of the trait, and from that to gain insight into the mechanisms of the

disease. Some researchers have conjectured that future advances in understanding will come from finding and examining rare variants associated with a strong effect on disease incidence (McClellan and King 2010). This conjecture assumes heterogeneity in the genetic factors underlying the medical traits (see also Ioannidis et al. 2007). The possibility of heterogeneity in the environmental factors as well (Taylor 2014a, 19ff) makes even more tenuous the heuristic connection from type A analyses (estimating heritability etc.) to type B analysis of measured genetic and environmental factors (see also Appendix 3). Moreover, the results of the GWA studies might make a proponent of type B analysis less optimistic than a decade ago about identifying associations of a *single* genetic factor and a *single* environmental factor.

The connection between the first two and the third kind of nature-nurture science—between group-average differences—is not addressed much in *Beyond Versus*. A relevant conjecture I have is that, just as there was a manipulable level of a measured environmental factor in Fisher's and Hogben's type C data analyses, the components of variation derived from type A analysis have been imagined by researchers (e.g., in debates about heritability of IQ test scores, p. 46ff) to correspond to measurable, albeit yet-to-be-identified genetic and environmental factors. It then seemed plausible that the same kinds of factors underlying variation within groups might be associated with the variation between groups (strictly, to the difference between the averages for the groups). Another, more sociological, conjecture is that genomics allows people to posit a hereditary basis for traits of medical or social significance, such as intelligence, and this bolsters and is bolstered by the power of selective breeding in agriculture and a persistent or revived eugenic ideal of improving society by eliminating the defective biology of individuals. (Appendix 4 provides perspective on this nexus by reviewing what actions are actually possible based on type A associations.)

The connection to the fourth kind of nature-nurture science—fixity versus flexibility of traits—is also not addressed much in *Beyond Versus*. However, one relevant body of research is that of Kendler and colleagues, who have examined incidence of depression in relation to a wealth of measured factors over the life course as well as a factor derived from the relatedness of the individuals, which they label "genetic risk." In Kendler et al. (2002), for example, data on over 1,900 twins are used

to fit the incidence of major depression in women to a model that accounts for 52% of the variance in the trait. The model is static, but it is structured to include connections from earlier-in-life to later-in-life factors, e.g., from risk by relatedness (genetic risk) to neuroticism to low self-esteem to low education through stressful life event to major depression, and thus provides a picture of development that is rich and plausible. (The "Structural Equation Modeling" approach used works as if the value of the trait in model B were, in turn, a factor on the right hand side of a second model B, and so on.) Interestingly, many of the factors are conceivably modifiable (e.g., women with low self-esteem could receive counseling), but no therapeutic or policy interventions are included in the factors examined even though only one of the many factors included in the implied picture of the development of depression fits firmly on the nature side. In any case, associations with the interaction contributions in type A, B, or C analysis cannot be decisive. The relevance of these associations depends, as mentioned under Move 2, on showing that the full range of locations or environmental factors has been included in the observations. Turkheimer at al. (2003) shows, for example, that heritability of IQ test scores, typically stated as being around 60%, is almost zero in families of low socioeconomic status.

In summary, the four different kinds of nature-nurture science are not speaking to the same issues. The connection often implied between type A and type B analysis is not warranted except as a heuristic that is tenuous anyway; the tenuous heuristic is no basis for assuming that the same kinds of factors underlying variation within groups are associated with the difference between the averages for the groups; and models of multiple measured genetic and environmental factors have yet to progress to a place where they can speak to developmental fixity versus flexibility.

Another alternative, in closing, to Move 1, Dividing sociological and philosophical inquiry

This essay has suggested alternatives to seven conceptual moves that are key to Tabery's account. However, with respect to Move 1, the essay has, for the most part, joined him on the philosophical side. How, it might be asked, is the conceptual

clarification we have both pursued supposed to make a difference? To the extent that conceptual clarifiers envisage there to be scientists in our audience, are we saying or implying that we make systematic and clear what you had not—or more systematic and more clear? Or, we are endorsing researcher A over researcher B, or we can extend researcher A's thinking? Whichever of these messages fits best, it seems that we want researchers to see that they have overlooked some things and decide to modify the science with the goal of improving their accounts of the phenomena that constitute reality. Yet, as sociology and history of science remind us, critique—which includes conceptual clarification—is rarely decisive in shifting science. The production of scientific knowledge also involves many, diverse practical considerations as researchers employ equipment, experimental protocols, citations, the support of colleagues, the reputations of laboratories, metaphors, rhetorical devices, publicity, popular debates, funding, and so on (Latour 1987; Law 1987). Researchers linking such heterogeneous resources are also traversing different domains of social action—*social worlds*—to which they contribute to continuity as well as change (Clarke and Fujimura 1992). (The diverse social worlds that intersect in the nature-nurture sciences are readily seen by searching "nature nurture" in google books; see Appendix 1.)

Even when conceptual clarification is directed, more modestly, at commentators on nature-nurture science, including philosophers of science, how much difference can be made by a book or a book review is an open question. The answer would depend on how much the attention, topics, funding, metaphors, and so on of the commentators draw from the many and various currents of the genomic era (Taylor 2014b). To invoke this sociological sphere is not to counsel despair for those of us with a penchant for conceptual clarification. Rather, we can view the sociological embeddedness of our audiences and of ourselves as an invitation to extend the necessarily partial contributions we make to modifying scientific knowledge. Examination of conceptual developments within the sciences can lead us into interpretive questions about the social influences shaping scientists' work or its application, which, in turn, can lead to new questions and awareness of alternative approaches in those sciences (Taylor 2014a, 42ff). In this spirit, we could delve into the cross-reinforcement of distinct nature-nurture sciences or the conflation of different meanings of interaction. We could

puzzle over how associations exposed by statistical analysis are so readily translated into proposals for action. And so on. Many interesting inquiries remain for conceptual clarifiers and other commentators who want to help science and society—indeed, science-in-society—move beyond nature versus nurture.

Works Cited

Byth, D. E., R. L. Eisemann, et al. (1976). "Two-way pattern analysis of a large data set to evaluate genotypic adaptation." Heredity 37(2): 215-230.

C Reactive Protein Coronary Heart Disease Genetics Collaboration (2011). "Association between C reactive protein and coronary heart disease: Mendelian randomisation analysis based on individual participant data." British Medical Journal 342(15 February): d548.

Caspi, A., J. McClay, et al. (2002). "Role of genotype in the cycle of violence in maltreated children." Science 297: 851-854.

Caspi, A., K. Sugden, et al. (2003). "Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene." Science 301: 386-389.

Clarke, A. and J. Fujimura (1992). "What tools? Which jobs? Why right?" In A. Clarke and J. Fujimura (Eds.), The right tools for the job: at work in twentieth-century life sciences (pp. 3-44). Princeton: Princeton University Press.

Davey-Smith, G. and S. Ebrahim (2007). "Mendelian randomization: Genetic variants as instruments for strengthening causal influences in observational studies." In M. Weinstein, J. W. Vaupel and K. W. Wachter (Eds.), Biosocial Surveys (pp. 336-366). Washington, DC: National Academies Press

Hargreave, M., A. Jensen, et al. (2013). "Fertility treatment and childhood cancer risk: a systematic meta-analysis." Fertility and Sterility 100(1): 150-161.

Ioannidis, J. P. A., N. A. Patsopoulos, et al. (2007). "Heterogeneity in meta-analyses of genome-wide association investigations." PLoS ONE 2(9): e841.

Jensen, A. R. (1969). "How much can we boost IQ and scholastic achievement?" Harvard Educational Review 39: 1-123.

Jensen, A. R. (1970). "Race and the genetics of intelligence: A reply to Lewontin." Bulletin of the Atomic Scientists 26: 17-23.

Kendler, K. S., C. O. Gardner, et al. (2002). "Towards a comprehensive developmental model for major depression in women." American Journal of Psychiatry 159: 1133-1145.

Latour, B. (1987). Science in action: how to follow scientists and engineers through society. Milton Keynes: Open University Press.

Law, J. (1987). "Technology and heterogeneous engineering: The case of Portuguese expansion." In W. E. Bijker, T. P. Hughes and T. J. Pinch (Eds.), The social construction of technological systems: new directions in the sociology and history of technology (pp. 111-134). Cambridge, MA: MIT Press.

Lewontin, R. C. (1970). "Further remarks on race and the genetics of intelligence." Bulletin of the Atomic Scientists: 23-25.

Lewontin, R. C. (1974). "The analysis of variance and the analysis of causes." American Journal of Human Genetics 26: 400-411.

Lewontin, R. C. (1982). Human diversity. New York: Freeman Press.

McCarthy, M. I., G. R. Abecasis, et al. (2008). "Genome-wide association studies for complex traits: consensus, uncertainty and challenges." Nature Reviews Genetics 9(May): 356-369.

McClellan, J. and M.-C. King (2010). "Genetic heterogeneity in human disease." Cell 141: 210-217.

Mustillo, S. A., N. Krieger, et al. (2004). "The association of self-reported experiences of racial discrimination with black/white differences in preterm delivery and low birth weight: the CARDIA study." American Journal Public Health 94: 2125-2131.

Paul, D. B. and J. P. Brosco (2013). The PKU paradox: a short history of a genetic disease. Baltimore: Johns Hopkins University.

Plomin, R. (1999). "Genetics and general cognitive ability." *Nature* 402: C25-C29.

Plomin, R., J. C. Defries, et al. (1997). Behavioral genetics. New York: Freeman.

Plomin, R., J. C. DeFries, et al. (1977). "Genotype-environment interaction correlation in analysis of human behavior." Psychological Bulletin 84(2): 309-322.

Rader, K. A. (2004). Making mice: standardizing animals for American biomedical research, 1900-1955. Princeton: Princeton University Press.

Ridker, P. M., J. E. Buring, et al. (2007). "Development and Validation of Improved Algorithms for the Assessment of Global Cardiovascular Risk in Women: The Reynolds Risk

Score." Journal of the American Medical Association 297: 611-619.

Taylor, P. J. (2014a). Nature-nurture? No: Moving the sciences of variation and heredity beyond the gaps. Arlington, MA: The Pumping Station.

Taylor, P. J. (2014b). "Whys to look for genes: pros and complications." Retrieved 2 January 2015, from http://whystolookforgenes.wordpress.com/about/.

Turkheimer, E., A. Haley, et al. (2003). "Socioeconomic Status Modifies Heritability of IQ in Young Children." Psychological Science 16(6): 623-628.

Appendix 1. A review of books that clarify or recast nature-nurture issues

The list of books below is a selection derived from searching "nature nurture" in google books (see http://bit.ly/NvNbooks). The table to follow summarizes positions that run through more than one book.

Number	Position	Contrasting position presented in the body of the essay
1.	Heritability is a measure of the effects of genes and the remainder of the variation is the effect of the environment.	Not correct. Type A analyses (estimating heritability etc.) have a tenuous heuristic connection with type B analysis of measured genetic and environmental factors (see also Appendix 3).
2.	The contribution of the environment can be partitioned into shared and non-shared components. (The non-shared eclipses the shared.)	No position stated in the essay. However, just as in 1, the position implies a connection between type A and B analysis and is thus questionable (Taylor 2014a, 115-116).
3.	Nature-nurture debate refers to relative strength of genetic and environmental influences in the development of traits of an individual, in variation across a population, and in differences between groups.	Position 3 combines four kinds of nature-nurture science, which have no clear conceptual or empirical connection between them.
4.	Traits are caused by complex interactions between genes and the environment at every stage of biological and psychological development.	In a dynamic sense of interaction, yes. The nature-nurture sciences reviewed in this essay concern, however, the analyses of quantitative data that do not directly address development and in which interaction is a statistical concept. Investigations that extend the analysis of observational data are needed for researchers to piece together a picture of the processes of development of a trait (and, on that basis, speak to the fixity versus flexibility of traits).

5.	In analyses of observational data categorized by some genetic factors and environmental factors contributions* provide insight about intervention to alter the trait in question. [*meaning as given in this essay]	Contributions in any analysis of variance (types A-C): a) are conditional on the full set of individuals (or varieties) and situations (or locations) where they are observed; and b) are not necessarily modifiable to reproduce that variation.

Book	**Position in relation to issues raised in the body of the essay**	**Position in relation to other issues**
Fausto-Sterling, A. (1985). Myths of Gender. New York: Basic Books.	4.	Genetic does mean unchangeable. In particular, children show flexibility in gender self-concept. Views of sexual and gender development are biased towards male development.
Flynn, J. R. (2012). Are We Getting Smarter? Rising IQ in the Twenty-First Century. Cambridge: Cambridge University Press.		Focuses on marked differences between generations in average test scores, with some attention to differences between racial groups and gender. Offers explanation of the former (rejecting hypotheses that genetics, in the form of outbreeding, could be involved). Issue not settled whether racial differences are genetic or environmental in origin.
Goldhaber, D. (2012). The Nature-Nurture Debates. Cambridge: Cambridge University Press.	4, but presents 1, 2, and 3 without disputing them on their own terms.	Nature-nurture science (1-3) does not provide insight into the development of traits of an individual or the influence of development on evolution.
Harris, J. R. (2009). The nurture assumption: Why Children Turn Out the Way They Do. New York: Free Press (2nd Ed.)	1, 2, 3 + the environment is shaped by the child's genes, not by shared family upbringing.	The child's interaction in peer groups is the primary and natural determinant of their personality.

Kaplan, G. and L. J. Rogers (2003). Gene Worship: Moving Beyond the Nature/Nurture Debate over Genes, Brain, and Gender. New York: Other.	Addresses separately all four kinds of nature-nurture science.	Focusing especially on gender and sex development, modifiable interactions occur at every stage of development.
Keller, E. F. (2010). The Mirage of a Space between Nature and Nurture. Durham, NC: Duke University Press.	1 + notes slippage between heritability and heritable.	More attention is needed to what genes do during development.
Lewontin, R. C., S. Rose, et al. (1984). Not in Our Genes: Biology, Ideology and Human Nature. New York: Pantheon.	Careful in definition and interpretation of heritability (but not explicit about the contrasting position to 1 presented in the body of the essay).	High heritability does mean unchangeable.
Lewontin, R. C. (2000). It Ain't Necessarily So: The Dream of the Human Genome and Other Illusions. New York: New York Review of Books (esp. Chapter 1)	1.	Heritability provides no information about changeability (see alternative to 5).
Longino, H. (2013). Studying human behavior: How scientists investigate aggression and sexuality. Chicago: University of Chicago Press.	1, 2, 3 + heritable and heritability treated as synonyms.	Pluralism: Each method of studying behavior is best seen as a partial view, not a competing view to other methods, such as molecular behavioral genetics.
Moore, D. S. (2001). The Dependent Gene: The Fallacy of "Nature vs. Nurture". New York: W. H. Freeman.	4, but presents 1 & 3 without disputing them.	The focus of research should be on the development of traits of an individual (for which 1 & 3 are not helpful).

Pinker, S. (2002). The Blank Slate: The Modern Denial of Human Nature. New York: Viking.	1, 2, 3 + heritable and heritability treated as synonyms.	Opposition to hereditarian explanations follows from the theory that the mind is a blank slate, which distorts understandings of gender, upbringing, violence, and more.
Plomin, R. (1990). Nature and Nurture: An Introduction to Behavioral Genetics. Pacific Grove, CA: Brooks/Cole.	1, 2	
Ridley, M. (2003). Nature Via Nurture : Genes, Experience, and What Makes Us Human. London: Fourth Estate.	1, 2, 3, 4, 5 (with reservations about the significance of heritability and non-shared environmental components)	The action of genes (nature) is influenced by experience (nurture). The effect of that experience or environment varies with the genes an organism has.
Rutter, M. (2006). Genes and Behavior: Nature-Nurture Interplay Explained. Malden, MA: Blackwell.	1, 2, 5 (except questions that non-shared eclipses the shared).	Special attention given to interaction between measured genetic and measured environmental factors.
Tomasello, M. and D. I. Slobin (2005). Beyond Nature-Nurture: Essays in Honor of Elizabeth Bates. Mahwah, NJ: Lawrence Erlbaum.		In the development of language, knowledge comes from the "interaction between genes, bodies, and environments, unfolding overtime… and the ways in which structures can arise without being prespecified."

Appendix 2. Three types of model for statistical analysis of observational and experimental data

$y_{ijk} = m + v_i + l_j + vl_{ij} + r_{ijk}$ (A)

where y_{ijk} denotes the observed value of the trait y for the i^{th} variety* in the j^{th} location* and k^{th} replication, which is modeled as a sum of the following variables:

m for a base level for the trait;

v_i for the contribution of the i^{th} variety;

l_j for the contribution of the j^{th} location;

vl_{ij} for the additional contribution from the i,j variety-location combination not already given by the preceding two contributions; and

r_{ijk} for the residual contribution.

$y_{ijk} = m_B + \alpha g_i + \beta e_j + \gamma g_i e_j + r_{ijk}$ (B)

where y_{ijk} is observed value of the trait y for the i^{th} measured genetic factor under the j^{th} measured environmental factor and k^{th} replication, which is modeled as a sum of the following measured factors multiplied by coefficients a, b, g:

m_B for a base level for the trait;

αg_i for the contribution of measured genetic factor i;

βe_j for the contribution of measured environmental factor j;

$\gamma g_i e_j$ for the additional contribution from the i,j gene-environment combination not already given by the preceding two contributions; and

r_{ijk} for the residual contribution.

$y_{ijk} = m_C + v_i + \beta e_j + \gamma_i e_j + r_{ijk}$ (C)

where y_{ijk} is observed value of the trait y for the i^{th} variety under the j^{th} experimentally manipulated environmental factor and k^{th} replication, which is modeled as a sum of the following variables and measured factors multiplied by coefficients β, γ_i:

m_C for a base level for the trait;

v_i for the contribution of the i^{th} variety;

βe_j for the contribution of experimentally manipulated environmental factor j;

$\gamma_i e_j$ for the additional contribution from the i,j variety-environment combination not already given by the preceding two contributions; and

r_{ijk} for the residual contribution.

* Varieties are often called *genotypes* despite the lack of knowledge of the measurable genetic factors that underlie the relatedness of individuals in a variety/genotype. Similarly locations are called environments without knowledge of the environmental factors present in each location/environment. In any case, in the equation for type A, to use the symbols g_i and e_j for genotype and environment would invite confusion with the conceptually and empirically distinct terms in the equation for type B.

Appendix 3. Heritability of a trait does not measure the contributions of genetic differences to observed differences among individuals

1. Readers with a technical knowledge of heritability may note that, when estimating heritability from datasets in which varieties have varying degrees of genealogical relatedness (e.g., identical or monozygotic twins versus fraternal or dizygotic twins), models often refer to *theoretical genes* that each add a small contribution to the trait. However, analyses built around these models are of observations of *traits*, so there must be alternative formulations making no reference to genes (Taylor 2014a, 55-76).

2. Consider one way to estimate heritability for a human trait, namely, comparing the similarity of identical twins, who share all their genes, with the similarity of fraternal twins, who share a smaller fraction; in both cases, the twins are raised together. Even if the similarity between twins or a set of close relatives is associated with the similarity of yet-to-be-identified genetic factors, the factors may not be the same from one set of relatives to the next, or from one location to the next. In other words, the underlying factors may be *heterogeneous* (Taylor 2014a, 19).

3. The possibility of underlying heterogeneity disturbs any intuition that a measurable genetic factor (or composite of factors) runs through the differences among variety means. We would not assume such a genetic gradient exists if the varieties were from different species or taxonomic classes. Yet the partitioning of variation involved in estimation of heritability and other components could, in principle, be undertaken even if varieties were not from the same species (Taylor 2014a, 28ff).

4. Even if there were such a gradient, it is difficult to move from type A variation-partitioning to hypotheses about underlying measurable factors even in the ideal case of such analysis, namely, the full agricultural evaluation trial (see Appendix 4). For the analysis of human observations, where a variety (or genotype) is replicated at most two times in at most two locations (environments), it is not possible to group similar varieties and locations and, on that basis, generate hypotheses about underlying factors (Taylor 2014a, 30ff).

5. To add further ambiguity, in recent years the term heritability has begun to be used to refer to the fraction of variation in a trait associated with variation in Single-Nucleotide Polymorphisms (SNPs) as examined by Genome-Wide Association (GWA) studies. There is no conceptual connection between this new heritability and the classical concept (Taylor 2014a, 124-5).

6. The slippage between, or conflation of, type A and B analysis is fostered by ambiguous terms. In agricultural trials, varieties have often been called genotypes even though no claim is made that a pair of alleles—the strict meaning of genotype—defines the variety. Variation among the variety contributions in model A is then variation associated with difference between genotype means (or "genotypic values"), shortened to genotypic variance (*variance* being the technical statistical measure of variation) and then, unfortunately given its ambiguity, to genetic variance. Similarly, variation among location means is often referred to as environmental variance. Yet no genetic or environmental factors are used in the making of type A analyses; genetic variance does not refer to variation among measurable genetic factors.

Appendix 4. Actions based on type A analysis, including assessment of the degree of interaction, and to investigation of mechanisms

The distinction between type A and B analysis is illuminated by examining the possibilities and limitations of type A analysis with respect to actions based on the analysis, including assessment of the degree of interaction, and to investigation of mechanisms.

Consider, as an example of data and analysis of type A, the case of an agricultural evaluation trial where it is possible to observe a trait, say, yield, in a set of plant varieties in each of a set of locations, and to raise replicates for each variety-location combination (e.g., Byth et al. 1976). The variety contributions (see Figure 1 and Appendix 2) can be simply given by subtracting the overall mean for the data from the means of each variety when averaged over all the locations and replications in which it is grown. Similarly, for the location and variety-location interaction contributions. If the variation of the variety means is significant and we can imagine growing in the same locations the varieties that have the best mean yields, then we would expect to improve yields overall. Secondarily, plant breeders may cross the best varieties overall and expect yield improvement in proportion to how much of the variation is associated with the variety means. If some of the crosses turn out not to yield well, they can be discarded, while only those that yielded well get used.

Now, if the variety-location interaction contribution is also significant—which is typically the case in large crop evaluation trials (e.g., Byth et al. 1976)—then one variety may be highest yielding in one location but not in another—or, at least, the difference between any two varieties may change substantially from location to location (Taylor 2014a, 55). Interaction in this type A sense makes it difficult for agricultural researchers to provide a single recommendation to farmers on which variety to plant. Recommendations about what to grow need, instead, to be tailored to some subset of the locations.

The last two paragraphs indicate that crop recommendations and plant breeding decisions can proceed without an understanding of mechanisms. Indeed, the variables associated with the partitioned variation are not modifiable things. However, while such actions are being taken, researchers can also go on to investigate further to expose

what it is happening in the locations with the best yields. As described in Taylor (2014a, 95ff),

> varieties can be grouped by similarity in responses across all locations using techniques of cluster analysis (Byth et al. 1976). Similarly, locations can be grouped by similarity in responses elicited from varieties grown across those locations. Varieties in any resulting group tend to be above average for a location in the same locations and below average in the same location. The wider the range of locations in the measurements on which the grouping is based, the more likely it is that the ups and downs shared by varieties in a group are produced by the same conjunctions of measurable factors… For example, imagine a group of plant varieties that originated from particular parental or ancestral stock that is more susceptible to plant rusts (a form of parasitic fungi), and that these varieties had a poor yield in locations where rainfall occurred in concentrated periods on poorly drained soils. The obvious hypothesis about genetic factors modulated by environmental factors is that these varieties share genes from the parental stock that are related to rust susceptibility and this susceptibility is evident in the measurements of yield in locations where the rainfall pattern enhances rusts.

On the basis of such a hypothesis, plant breeders might cross the rust-susceptible varieties that yield well in other locations with rust-resistant varieties and look for progeny that are resistant and yield well in the rust-promoting locations. Alternatively, the original varieties might be evaluated in the same sites but with enhanced soil drainage practices. Success or failure in exposing mechanisms at the level of, for example, parental stock used in crosses or soil drainage will affect whether the researchers see any need to delve further into the genetic and environmental factors that influence the yields shown by the varieties in the various locations where they are grown. In other words, whether knowledge about mechanisms is needed for progress in agricultural science based on type A analyses *all depends*.

Now let us contrast the situation when the observations are of traits in human populations. First, the option of selective breeding and discarding crosses that turn out not to “yield” well is not available. Nor is it possible to reduce the interaction

contribution by grouping of varieties and subsequently forming hypotheses based on those groups. In short, type A analysis of human data does not provide a basis for actions based on the analysis or investigation of mechanisms. Regarding the assessment of the degree of interaction, it is possible, given the appropriate data sets, to separate the variety contributions from those of the variety-location (genotype-environment) combinations. The data needed is, for the one population, of twins raised apart, twins raised in the same family, and unrelated individuals raised in the same family (Taylor 2014a, 62). Such data sets are, however, rare. Plomin et al. (1977) is often cited in the context of claims that such interaction variation is not significant for humans, but this work considers as a proxy for variety-location interaction a quantity derived from a type B analysis of, for example, data on educational attainment, in which the measured factors are the average for biological parents and the average for adoptive parents. Tabery (152ff) reviews the evidence for low values of such proxy measures, but how well they reflect the actual type A variety-location interaction is hard to assess in the absence of studies for a range of human traits in which the classes of data are collected that allow the separation of variety from variety-location contributions. It is necessary to show that the latter contributions are negligible in order to be sure that reported heritability estimates for human traits capture only the differences among variety (genotype) contributions. Ditto, to interpret the trend that Plomin (1999, C26) and others have noted for heritability estimates to increase over people's lifetimes. It could be that the interaction contribution, subsumed in estimates labeled as heritability, is increasing over time. For this reason, over and above the points made in Appendix 2, interpreting this trend as evidence that genetic differences come to eclipse environmental differences (Plomin 1999, C26) is not warranted (Taylor 2014a, 73-74).