Biol Philos (2017) 32:25–49
DOI 10.1007/s10539-016-9535-1

Heritability and causal reasoning

Kate E. Lynch[1]

Received: 14 April 2016 / Accepted: 1 August 2016 / Published online: 13 August 2016

Abstract Gene–environment (G–E) covariance is the phenomenon whereby genetic differences bias variation in developmental environment, and is particularly problematic for assigning genetic and environmental causation in a heritability analysis. The interpretation of these cases has differed amongst biologists and philosophers, leading some to reject the utility of heritability estimates altogether. This paper examines the factors that influence causal reasoning when G–E covariance is present, leading to interpretive disagreement between scholars. It argues that the causal intuitions elicited are influenced by concepts of agency and blame-worthiness, and are intimately tied with the conceptual understanding of the phenotype under investigation. By considering a phenotype-specific approach, I provide an account as to why causal ascriptions can differ depending on the interpreter. Phenotypes like intelligence, which have been the primary focus of this debate, are more likely to spark disagreement for the interpretation of G–E covariance cases because the concept and ideas about its 'normal development' relatively ill-defined and are a subject of debate. I contend that philosophical disagreement about causal attributions in G–E covariance cases are in essence disagreements regarding how a phenotype should be defined and understood. This moves the debate from one of an ontological flavour concerning objective causal claims, to one concerning the conceptual, normative and semantic dependencies.

Keywords Heritability · Causation · Genes · Environment · Phenotype

✉ Kate E. Lynch
kathleen.lynch@mq.edu.au

[1] Department of Biological Sciences, Macquarie University, Room 209, E8B, North Ryde 2109, Australia

Introduction

Imagine a future in which a particular gene variant for obesity is isolated. Individuals with this variant are 50 times more likely to be obese than those with a different variant. These results lead to an acceptance in the community that obesity is a genetically caused condition, alleviating societal stigma and influencing its disease-classification. Researchers invest in targeting gene expression as a medical intervention and their research uncovers the biological pathways associated with the variant. Unexpectedly, it is discovered that individuals with this gene variant are *not* predisposed to deposit fat or utilise energy in different ways to non-obese individuals. Instead, the difference between those with the gene variant and those without is that those with the 'obesity gene' self-report that they enjoy higher calorie foods far more than those without the same genetic background, and as a result seek out and consume a larger amount of calories on average than others. Under this account, is obesity still a genetically caused condition?

In this case genetic differences led to differences in environmental input between two genotypic groups. It is undisputed that genes and the environment are both causally essential for the development of phenotypes, acknowledging their inescapable interaction during development has been termed the interactionist 'credo' (Kitcher 2001), and this widely accepted platitude has been used to advocate dissolution of the nature–nurture debate. However, the false dichotomy between genes and environment represents a straw man for a substantial part of the discussion concerning nature and nurture. As the example above intended to show, more complicated causal relationships between genes and the environment can muddy the dialectical waters. Heritability[1] is a statistical estimation of the relative causal contribution of genetic variation (V_G)[2] [as opposed to environmental variation (V_E)] to phenotypic variation (V_P) within a population (Eq. 1). This statistic relies on the assumption that V_G and V_E act additively, so that is there is no interaction or correlation between the two terms (Eq. 2). Under this framework it is an empirical matter as to how much genetic or environmental differences contribute, validating the persistence of a nature–nurture discussion.

$$H^2 = \frac{V_G}{V_P} \tag{1}$$

$$V_P = V_G + V_E \tag{2}$$

In simple cases, a high heritability estimate is meant to correspond to some notion of genetic causation or determination of phenotypic variation (Block 1995, 116; Fisher 1918, 399; Sesardic 2005, 22). However, if there is a statistical relationship between V_G and V_E, the additivity of this model breaks down, and

[1] My focus in this paper is broad-sense heritability, which is the primary interest of behavioural geneticists (see Sesardic 2005, 21). However, many ideas in this manuscript can also be applied to the narrow-sense heritability (H^2) concept used in quantitative and evolutionary genetics.

[2] Where variation is represented by the statistical concept of variance. The V_G, V_E and V_P terms all denote variance.

along with it interpretations of genetic (and environmental) causality. There are two major ways in which additivity can be violated in a heritability model. The first is gene–environment interaction, which occurs when the effect of a change in value of one variable (such as V_G) varies depending on the values of the second variable (such as V_E). For example, environmental exposure to benzene is significantly associated with shorter gestation periods in pregnant women who possess a particular variant of the *CYP1A1* gene, whereas no such association exists in non-carriers (Wang et al. 2000). Thus the way in which the environment affects a phenotype (gestation period) is dependent upon the genetic background of the individual, and vice versa. This phenomenon complicates the partitioning of genetic and environmental variation as causes of phenotypic differences, as the value of one variable depends upon the other.

Once dismissed as a 'purely academic problem' (Fisher as cited in Tabery 2014, 33) gene–environment interaction has gained increasing attention in the field of behavioural genetics (Falconer and MacKay 1996; Plomin et al. 2008), and the philosophy of biology (Tabery 2014). The practical implications of gene–environment interaction are now widely accepted, and it is routinely factored into heritability estimates as an additional variable (V_{GxE}) using techniques such as analysis of variance (Falconer and MacKay 1996). Less attention has been paid to the second non-additive factor: gene–environment covariance (henceforth G–E covariance, also sometimes referred to as gene–environment correlation and rGE). G–E covariance occurs when there is a statistical association between genotypes and environments in a population, which is generally due to a causal relationship between individuals' genotypes and their environments. In the opening paragraph of this paper, an example case of G–E covariance was presented illustrating that the apportionment of nature (V_G) and nurture (V_E) in these cases is far from straightforward. When non-additivity is present, the heritability formula can be expanded to:

$$V_P = V_G + V_E + V_{GxE} + COV_{GE} \quad (3)$$

Opinions vary on how to interpret this phenomenon. Some (e.g. Block 1995; Block and Dworkin 1976; Gibbard 2001; Sober 2001) conclude that heritability estimates from studies with G–E covariance present do not accord with common-sense ascriptions of 'genetic causation', and as such reject the inclusion of any covarying V_G in the heritability statistic. Instead, they attribute the resulting variance as stemming from the environment (V_E), or believe that it should be partitioned into an additional variable (COV_{GE}). Others (e.g. Eaves et al. 1977; Roberts 1967; Sesardic 2005) argue for the inclusion of covarying V_G towards H^2, however only in certain cases. This presents a problem for the analysis of heritability, as there no consensus as to whether or not phenotypic variation in these situations has a genetic or environmental causal origin.

At present, it is unclear (1) exactly why theorists disagree about the inclusion or exclusion of covarying V_G to H^2, and (2) why some theorists opt to include G–E covariance in some circumstances and not in others. This paper will provide an account of these interpretive differences by arguing that non-causal factors which

contribute to causal reasoning processes impinge upon an assessment of G–E covariance cases. I draw on evidence from the literature on causal attribution which suggests that agency and blame-ascription play a role in the causal assignment made to variables across causally identical scenarios. These factors provide an account of why two types of G–E covariance (active and reactive) are often interpreted differently, despite their homologous causal structure. However, these factors alone are not sufficient to account for the differences in interpretation that occur. In order to completely explain the discrepancy in interpretations in the literature I appeal to the conceptual influence of the phenotype under study. I will argue that differences in phenotypic concepts and ideas about normal development are the final piece of the puzzle for accounting for interpretive differences in G–E covariance cases, particularly within active cases.

Gene–environment covariance

Passive gene–environment covariance

When parents are biologically related to their offspring they bestow both a developmental environment and their genes, which can lead to passive G–E covariance (Plomin et al. 1977). For example, intelligent parents tend to contribute both a genetic endowment and a scaffolded environment which allows for better development of intelligence in their children. Assuming some genetic basis for intelligence, a child's genotype will be correlated with their developmental environment when their parent's genetically influenced intelligence shapes that environment. As such, a covariance between children's genotypes and environments emerges within a population. It is termed passive G–E covariance because the shaping of the child's environment is in no way due to the consequences of his or her own genotype, but is instead due to the actions and causal influences of the genotype of their parents.

Block and Dworkin (1976, 480) refer to passive G–E covariance as providing either a 'double advantage' or a 'double disadvantage', where children inherit either higher genotypic intelligence and a better environment, or a lower than average intelligence and less stimulating environment. This implies that passive G–E covariance will always increase heritability, as it magnifies the phenotypic differences between genotypic groups, inflating the impact of V_G on V_P. However, passive G–E covariance[3] can principally also deflate H^2. Imagine a family with a genetic predisposition to weight gain of which the parents are aware, leading to an inherited environment where high calorie foods are restricted. In this situation genes for a high body mass are negatively correlated with an environment for high body mass, leading to a deflation of the effects of V_G on V_P.

The causal structure of passive G–E covariance is presented in Fig. 1. As heritability analyses concern the relative influence of the child's genotype (V_{Gc}) and

[3] This also applies the other two forms, reactive and active G–E covariance, discussed in "Active gene–environment covariance" section.

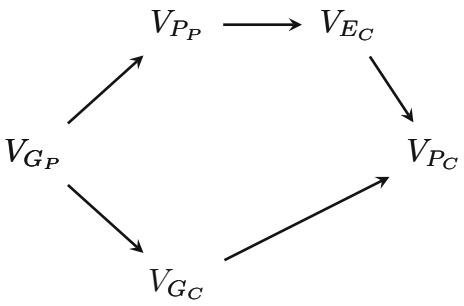

Fig. 1 The causal structure of passive G–E covariance. Variance in parental genotype (V_{G_P}) causes variance in children's genotype (V_{G_C}), as well as in the child's environment (V_{E_C}) via a genetically influenced parental phenotype (V_{P_P}). Both variance in the children's environment (V_{E_C}) and in their genotypes (V_{G_P}) causes variance in phenotype (V_{P_C}). However, the variance in children's genotype (V_{G_C}) does not causally influence V_{E_C}. Note that if the causal chain between V_{G_P} and V_{E_C} were broken, no passive G–E covariance would occur

the child's environment (V_{E_C}) on the child's phenotype (V_{P_C}), the covarying V_G and V_E can be considered non-causal. That is, there is no causal arrow connecting V_G [in this case (V_{G_C}) and V_E (V_{E_C})] other than the common cause of parental genotype (V_{G_P}). For this reason, passive G–E covariance is not extensively discussed as a criticism of heritability measures, for it can in principle be controlled for by methodological means,[4] as acknowledged by critics of heritability analyses (Block 1995; Jencks 1980; Sesardic 2005; Sober 2001). One such method of control is to compare the correlations of environmental conditions and children's phenotypes in adoptive and non-adoptive families (Plomin et al. 1977, 2008). This method has uncovered a significant passive G–E covariance for behavioural problems, temperament, and mental and language development (Plomin et al. 1985).

Reactive gene–environment covariance

Reactive G–E covariance occurs when a subject's developmental environment is imposed upon them, as a reaction to some expression of the subject's genetic background. Extreme versions of reactive G–E covariance are illustrated in a macabre series of thought experiments in which society singles out red-haired children and maltreats them, resulting in changes to other phenotypes, such as lowered IQ. In these examples it is assumed that hair colour has a genetic basis, and therefore one genotypic group (children with red hair) is subjected to a different environment (abuse) compared to another. This correlates genotypes and environments, and at a population level, V_G and V_E. Unlike the structure shown for passive G–E covariance cases, reactive scenarios display an indirect causal structure (Fig. 2, see also "Appendix"). V_G causes variation in some intermediate phenotype such as

[4] This method does not account for the passive covariance between genotypes and the maternal environment as all children (adopted or not) will experience a maternal environment that is correlated with their genotype. The maternal environment has been shown to significantly impact phenotypic variation in both humans and animals for a wide range of behaviours (Maestripieri and Mateo 2009) and has been suggested to contribute to IQ variation (Devlin et al. 1997).

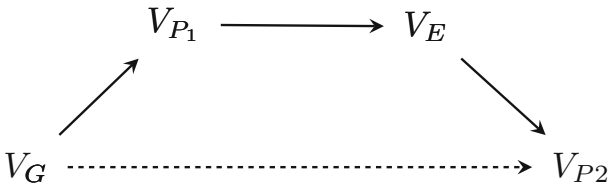

Fig. 2 The causal structure of active and reactive G–E Covariance. Variation in genotype (V_G) causes variation in a phenotype (V_{P1}), which causes variation in the environment which individuals experience (V_E). This in turn causes variation in a different phenotype (V_{P2}). V_G may also cause V_{P2} directly, as indicated by the dashed line. For more detail on this causal structure see "Appendix"

hair colour (V_{P1}), which causes variation in the environment experienced (V_E), which in turn causes variation in a secondary phenotype, such as IQ (V_{P2}). This may be made more fine-grained by including an additional variable after V_{P1} (V_X), representing differences in the way the foci individuals are treated by others, which is a more direct cause of variation in the environment experienced. In Fig. 2, V_X is encompassed as part of the more general variable V_E. Thus V_G causes V_{P2} indirectly via V_{P1} and V_E. Under an additive account of heritability (Eqs. 1, 2) the resulting phenotypic variation for IQ would be attributed to V_G because of this indirect causation via the covariant societal abuse variables (V_X and V_E).

These exaggerated examples demonstrate how prejudicial treatment based on genetically expressed variation in physiological phenotypes can contribute towards the apparent heritability of other phenotypes. This is one of the central criticisms of assertions that genetic variance causes racial differences in IQ (Jensen 1969; Herrnstein and Murray 1994). Critics have argued that rather than genetic differences causing racial differences in IQ is it is more likely that disadvantageous environments are correlated with populations of a particular racial background (Block 1995; Block and Dworkin 1976; Sober 2001).

The outcomes of sexual discrimination could be interpreted as heritable in a similar way. Gender is highly correlated with genotype, and this genetic difference typically causes differences in both physical appearance and gender identity, which affects the way individuals are treated in most societies. If a study were to use the sex chromosomes as a marker of genetic variance, then income, time spent doing house work, exposure to sexual assault, and the ability to drive in Saudi Arabia would all appear to be highly heritable. Yet the phenotypic variation in these, the racial, and the red-haired children examples are all intuitively considered to result from socialized environmental responses, and not genetic difference making. From these examples it is clear that reactive cases of gene–environment covariance are discordant with intuitive ideas about genetic causation. Based on these kinds examples Block (1995, 116) has claimed that heritability estimates are a misleading statistic which are in 'violent conflict' with common sense ideas about causation. Others have gone even further, claiming that heritability estimates are of no use at all for partitioning environmental and genetic causes of trait variation (Jencks 1980; Gibbard 2001; Sober 2001).

Active gene–environment covariance

While reactive G–E covariance has led some to reject the utility of H^2 altogether, less attention has been given to active forms. Active G–E covariance occurs when the environment is shaped by the actions of the individual possessing the covarying genotype, and these actions are at least in part genetically caused. As with the reactive cases, differences in environmental modification are associated with downstream phenotypic effects, and display an indirect causal structure (Fig. 2, see also "Appendix").

For example, imagine two genotypic groups of children where G1s have a small genetic advantage compared to G2s in terms of their intelligence (V_G). Because of this small and early advantage G1 children modify their environment in a way to intellectually stimulate themselves throughout development, by seeking out books, taking extra classes, and working on problems. This variation in environmental modifying behaviours can be thought of as a phenotype in itself, which differs between the two groups (V_{P1}), resulting in differences in experienced environments (V_E). These differences in environments are akin to variation in an 'extended phenotype', where genetic effects extend beyond the boundary of the individual (Dawkins 1982).[5] Consequently, G1 children reach a stage of measurable intellectual advantage, which is expressed as significant IQ[6] differences between the two genotypic groups (V_{P2}) (example adapted from Jencks et al. 1972). Thus 'small genetic differences may therefore end up producing big environmental differences' (Jencks et al. 1972, 110), which lead to large scale phenotypic differences.

In this example active G–E covariance inflates H^2, but as demonstrated in the passive case, it is possible for G–E covariance to have the opposite effect. Imagine a situation in which G1 children (those with a genetic advantage in intelligence) decide not to try as hard at school as they can get away with studying less while achieving similar results to other students. G2 children, lacking this genetic endowment, act in the opposite way, trying especially hard and seeking out extra stimulation and study materials to compensate for their shortcomings. Variation in IQ will be diminished between the groups as initial genotypic differences are compensated for by environmental selection. In this situation less of the overall V_P will be accounted for by V_G than if the G–E covariance had not occurred.[7] Thus G–E covariance should be generally thought of as having the potential to bias heritability estimates, and not necessarily inflate or deflate them.

It is not possible to separate active and reactive forms of G–E covariance experimentally. They are more difficult to estimate directly than passive forms, and as such are likely to go undetected to an even greater degree. However, there is good evidence that children actively shape their environment (Ambert 1997), and there is

[5] For a discussion on the way that the phenotype is defined in relation to G–E covariance see Lynch and Bourrat (forthcoming).

[6] It is assumed in this paper that IQ is a reasonably good proxy for general intelligence.

[7] A similar 'cancelling out' effect is also possible in the reactive framework. Parental or teacher encouragement could be bestowed upon children who are initially struggling, evening out the differences between genotype groups.

evidence that some self-mediated environmental alterations are based on genetic differences. For example differences in stressful life events as well as socioeconomic, educational, and occupational status have all been shown to be somewhat genetically mediated (Plomin et al. 2008; Rutter and Silberg 2002). Thus it is likely that each form of G–E covariance contributes significantly towards phenotypic variation. Yet the study of this factor in behavioural and quantitative genetics is vastly underrepresented. This appears to be, in part, due to a disagreement about the interpretation of these cases, which shall be discussed below.

Interpretations

Passive gene–environment covariance

As mentioned in "Passive gene–environment covariance" section, the general and uncontroversial interpretation of passive G–E covariance is that phenotypic variation arising from differences in inherited environments should be subsumed under V_E (Block 1995; Roberts 1967; Sesardic 2005; Sober 2001). As shown in Fig. 1 the covariance between V_G and V_E occurs because of the effects of the parental phenotype on V_E, and not the effects of V_G itself. Because of the general consensus in these cases, which coheres with the causal structure presented in Fig. 1, passive G–E covariance cases shall not be extensively discussed hereon.

Reactive gene–environment covariance

A similar consensus is reached for the interpretation of reactive cases, despite the fact that the majority of attention (and controversy) is paid to G–E covariance involves reactive examples. There have been no philosophers or biologists who defend the claim that the effects of reactive G–E covariance should be attributed as caused by V_G (Sesardic 2005; although see Lynch and Bourrat forthcoming).[8] Interpretative disagreement for reactive cases is more subtle. Fuller (1979, 427) makes explicit that the resulting variation in these cases should be considered as part of V_E, while others believe that these cases identify a source of variation that should be encompassed under a separate term (COV_{GE}), using an extended heritability Eq. (3) (Eaves et al. 1977; Jencks 1980; Loehlin and DeFries 1987).

Under this account, phenotypic variance arising from reactive G–E covariance does not contribute towards H^2, but is also not considered as stemming from environmental variation. Many though, do not specify how this source of variation should be treated, only that it should not be considered as part of genetic variance (Block and Dworkin 1976; Gibbard 2001; Sober 2001). While there is general

[8] Some have accused Roberts of defending the inclusion of reactive G–E covariance in H^2, due mostly to this quote: '…it matters not one whit whether the effects of the genes are mediated through the external environment or directly, through, say, the ribosomes' (Roberts 1967, 218). However, Sesardic (2003, 2005) points out that this allegation is misconceived, and Roberts in the above statement is referring to active G–E covariance.

agreement that the resulting V_P in reactive cases should not be ascribed to V_G, very little is said about *why* these cases should be interpreted in such a way.

Block (1995) and Block and Dworkin (1976) distinguish between 'direct heritability' in which the causal chain from gene to phenotype does not extend beyond the boundary of the organism and is mediated solely by "internal biochemical processes". 'Indirect heritability' on the other hand includes the effects of causes which may have originated within an individual, but then extend outside of the boundary of the organism. According to the authors only direct heritability cases are acceptable for use of the H^2 statistic—excluding both reactive and active cases of G–E covariance. A consequence of this view is that extended phenotypes are excluded completely from the study of heritability. This rules out the effects of genetic variance on any behavioural differences that are mediated by an intermediate environment, which excludes a large range of behaviours of interest to geneticists. In fact, if the direct/indirect distinction is taken seriously, then very few phenotypes fit the heritability criteria. This is because very few phenotypes arise out of causally closed and internally limited biochemical processes.

Block and Dworkin use the example of height as a phenotype that could be directly heritable. Under their account, differences in genes (V_G) cause different proteins (V_{P1}) which cause differences in the secretion of a pituitary hormone (V_{P2}) which affect height (V_{P3}). All of the steps from V_G to V_{P3} are internal biochemical processes, and it is presumed that variation in diet and other factors (V_E) influence height independently from this process. But this is not the case. While V_G does affect height in some 'direct internal sense' through pituitary hormones, it also affects things outside the organism—such as the nutrition that is received into the body. For instance, the *GHRL* gene produces ghrelin, a molecule that is secreted in the gastrointestinal tract and affects pituitary hormone secretions which regulate growth. But this protein also affects the hypothalamus, resulting in an appetite response which in turn affects how much nutrition the body receives via human action within their environment. The expression of ghrelin is also mediated by environmental cues, such as over or under nutrition. These cues are themselves affected by ghrelin expression via appetite, feeding back into the causal pathway involved (Burger and Berner 2014). Further, the *GHRL* gene is variable at a population level, and variation in this gene is thought to account for some of human height differences (Baessler et al. 2005).

Direct and indirectness aside, other justifications for excluding reactive G–E covariance tend to appeal solely to common sense and intuition. For example:

> Attributing redheads' illiteracy to their genes would probably strike most readers as absurd under these circumstances. (Jencks et al. 1972, 66–67)

Sesardic (2003, 2005) has carefully identified many of these appeals to intuitions, but does not question their validity as a basis for the interpretation of heritability statistics. Instead, he takes pains to demonstrate that in practice geneticists generally attribute reactive G–E covariance to non-genetic factors, cohering with genetic effects 'in our usual sense of the word' (Sesardic 2003, 1004) that are 'not really anomalous or aberrant' (Sesardic 2005, 104) by using their best common-sense causal intuitions. According to Sesardic then, scientists are interpreting reactive

cases in accordance with causal intuition most of the time, and as such there is no issue within the sciences regarding interpretation. However, he does concede:

> This is all admittedly pretty vague, and I am not sure how intuitions underlying our different approaches to these two kinds of cases [active versus reactive] should be refined further and made more precise. Fortunately this doesn't really matter, for I only want to claim that in dealing with G–E correlations, behaviour geneticists are by and large guided by the common-sense considerations about causality, with all their characteristic vagueness and ambiguities. (2003, 1012–1013)

My contention is it does matter, and the following sections are devoted to understanding the foundations of the 'common-sense' ascriptions of causality that are being used to assess G–E covariance cases. Before embarking on this, it is necessary to consider the less publicised, but much more contentious, active form of G–E covariance. This case is particularly important, as intuitions about 'common sense ascriptions' of causation in these situations appear to be divided within the scientific and philosophical community.

Active gene–environment covariance

Unlike the other two kinds of G–E covariance, there is widespread disagreement about how to interpret active cases. This is of particular interest because, as shown in Fig. 2, active and reactive cases have symmetrical causal structures. One might expect that causal similarity would entail similarity of interpretation, but the literature suggests otherwise indicating that additional non-causal factors are at play. Some argue that in cases of active G–E covariance differences in environmentally modifying behaviours that are caused by genetic differences should be considered as part of the differences in the phenotype that one is measuring (V_P) (Eaves et al. 1977; Jencks 1980; Jensen 1969; Roberts 1967). This is in line with Dawkins' (1982) extended phenotype approach. Under this interpretation the resulting phenotypic variance from active G–E covariance should be attributed to V_G as the cases simply reflect a particular expression of genetic variation, where covarying environmental causes are a natural extension of the phenotype, which '…present(s) no more of a dilemma than the observation that fast growing genotypes eat more' (Eaves et al. 1977, 9).

Others think that any resulting variation from active G–E covariance derives from a different source of variation (Block 1995; Block and Dworkin 1976; Gibbard 2001; Plomin et al. 1977; Sober 2001). As with the reactive cases, it is often not specified whether this means it should be ascribed to V_E, or an alternative variable (Cov_{GE}). Like in the reactive cases, the basis of this assessment is the causal intuitions elicited by active G–E covariance examples.

It should be noted that for all of these interpretations, disagreement tends to be understood in an ontological manner. Philosophers, psychologists and behaviour geneticists are debating which interpretation of these situations is the *correct* one, appealing to some ontological causal truth underlying these cases, as illustrated by Roberts (1967, 234): "The overriding concern at this stage is to avoid

environmental sources of covariance that would lead to the wrong answer by inflating the estimate of the heritability…". This is often then taken further to prescriptive claims about how these cases *should* be interpreted. I shall return to this issue briefly in "Conclusion" section.

A less ontologically grounded conjecture has been made by Sesardic (2005), who suggests that the common-sense interpretation of active G–E covariance cases will vary depending on the nature of the particular case, analogous to some forms of contextual causal reasoning (2005, 104). Thus:

> …active G–E covariance is occasionally subsumed under genetic variance… it is important to stress that this redistribution is not a necessary consequence of some esoteric methodology for calculating heritability. Rather, it is a practical decision primarily guided by an attempt to follow the commonsense way of apportioning causal responsibility… (2005, 104)

The following sections uncover what factors drive the common-sense causal ascriptions that seem to be accounting for interpretative differences between cases and among academics.

Factors in causal reasoning: agency and blame

"Interpretations" section demonstrated that there is a general consensus about how to interpret passive G–E covariance cases (as attributable to V_E) and reactive cases (*not* stemming from V_G), while the correct way to interpret active cases remains source of disagreement among scholars (V_G, V_E, or Cov_{GE}). "Gene–environment covariance" section demonstrated that both reactive and active G–E covariance cases have identical causal structures (Fig. 2). Thus the causal underpinnings of these cases are not sufficient to account for differences in how they are causally interpreted. Understanding the interpretation of these cases is twofold. First, one must account for the interpretive differences *between* active and reactive G–E covariance cases, despite their identical causal structures. Secondly, the differences amongst scholars *within* active G–E covariance interpretations must be accounted for.

To begin with the first project: accounting for differences between active and reactive G–E covariance, I turn to the literature on causal attribution. Studies have shown that non-causal factors can influence the way in which causation in a scenario is interpreted. One of these is the absence or presence of a human agent in the causal scenario. Alicke (1992) demonstrated this by presenting vignettes to respondents about the cause of a car accident which varied in the number of agents involved. In these scenarios, the driver, John, is speeding home and involved in an accident. In some versions of the event an additional agent was involved in causing the accident (another driver), and in others an inanimate object played the same causal role (either a tree branch or an oil spill). It is important to note that across the three cases the actions of John and the causal structures linking John to the car accident remained the same. Alicke (1992) found that the inclusion of an additional agent in the system led to a decrease in the causal attributions made to John compared to the

non-agent scenarios, concluding that the inclusion of another human agent influenced judgments of causation via a diffusion of causal responsibility.

Agency alone does not appear sufficient to account for these causal attributions. An example from Hart and Honore (1985) helps to illustrate why. Imagine a householder who is prudent in storing firewood in his cellar, only to facilitate a pyromaniac the opportunity to burn the place down. Here both the householder and the pyromaniac are agents, and both have causally contributed to the house burning. Yet it seems intuitive that the pyromaniac is causally responsible for the house-burning, and not the householder. This is due to the blame-worthiness of the pyromaniac: culpability also plays an important role in causal attribution.

This principle was tested in Alicke's study by varying John's motivation for speeding. In one version John sped so that he could hide an anniversary present he had bought for his parents before they got home. In the alternative scenario he was speeding home to hide some cocaine that he had left out so that his parents would not find it. Although the actions, causal structures, and the outcome of the crash were identical in each case, John was identified as being causally responsible for the car crash more often in the cocaine-hiding scenario than in the present-hiding one.

An alternative account of determining a salient cause from a set of contributing causes is that the factor(s) that exhibit a deviation from what is contrasted as a normal state of affairs are selected (Knobe 2006). Under this interpretation moral culpability can be viewed as a norm deviation—as blameworthy acts diverge from what is morally normal. Knobe and Fraser (2007) tested this possibility by using a different vignette in which two agents differed in their blame-worthiness, but neither behaved in a way that deviated from normal expectations. Results indicated that blame-worthiness does indeed play an independent role in causal attribution.

To return to active and reactive cases of G–E covariance, it appears that a salient difference between the two is the presence of another blame-worthy agent. In active G–E covariance cases an individual seeks out and modifies their own environment, requiring no other agent acting in the causal chain between genotype and phenotype. In reactive G–E covariance cases *another* individual imposes an environment upon the subject by depriving them of educational resources. To illustrate this further, take the following two scenarios:

1. G1 children have pale skin, and because of the reaction that society has to pale-skinned children they are attacked with hot rods and their skin is burnt whenever they go outside. As a result these children spend more time indoors away from their peers and develop poor social skills. G2 children, who have darker skin, are not subjected to burning attacks from outsiders, and so spend more time outside with other children, developing better social skills.

This is a case of *reactive* G–E covariance. Differences in genotype result in differences in skin colour, which are reacted to differently in the environment (one genotype group is burnt while the other is not). As a result their environments differ—G1 children spend less time outside, and this causes another phenotypic difference between the groups—variation in social skills. Like with most of the reactive G–E covariance cases it seems unintuitive to ascribe the variation in social

skills to genetic differences. Instead, one would point the finger at the environmental differences between the groups—the differences made by the actions of other agents. Now consider the following alteration:

2. G1 children have pale skin, and because of the way that the sun reacts to pale skinned children, their skin is burnt whenever they go outside. As a result these children spend more time indoors away from their peers and develop poor social skills. G2 children who have darker skin, are not subjected to burning from the sun, and so spend more time outside with other children, developing better social skills.

This is a case of *active* G–E covariance. Differences in genotype result in physiological differences in the children in terms of how their skin burns. As a result they make active modifications to their environmental experiences, and the outcome is that they differ phenotypically. G1 children spend less time outside compared to G2 children, causing another phenotypic difference between the groups: variation in social skills. The only point of difference between cases 1 and 2 is the presence of a blameworthy agent. In the reactive example (1) other individuals are responsible for the burning of the pale-skinned children. In the active example (2) it is the sun, which is not an agent, and cannot be assigned blame in any meaningful way.

So the key difference between active and reactive G–E covariance is not a causal one, but relies on the presence or absence of a blameworthy agent. This accounts for the first step in understanding interpretive differences in reactive and active cases. But it does not explain the whole picture. If agency and blameworthiness were the only relevant features impacting the causal interpretations for these cases, then we would expect that all active G–E covariance cases are interpreted in the same way all of the time. The section below will demonstrate why this is not the case.

The importance of the phenotype

In this section I will turn to another factor that I believe influences causal ascription in biological cases—the phenotype under investigation. I show how causal intuitions for active G–E covariance are intimately tied with the conceptual understanding of the phenotype in question. In particular, phenotypes like intelligence which relate to a relatively ill-defined and debated set of cognitive capacities are more likely to spark disagreement for the interpretation of active G–E covariance. This observation is especially pertinent as a large amount of heritability research has focussed on intelligence and related traits. A focus on these phenotypes is likely to have biased the interpretation of active G–E covariance more generally in behavioural genetics.

In order to demonstrate how phenotypes influence causal interpretations of heritability results, I will present examples in which no G–E covariance occurs, and ones in which active G–E covariance does occur, for four different phenotypes. Each phenotype used in this section has been estimated as heritable, and I will assume for the purposes of this discussion that those estimates are reliable. The

examples in which no G–E covariance occurs should be straight-forward in their interpretation, where V_P that is due to genotypic differences is ascribed to V_G, and counts towards H^2. Non-G–E covariance cases are uncontroversially interpreted in this way, no matter the phenotype. It is in the active G–E covariance cases that vary in their interpretation, and I shall show how the phenotype under investigation is responsible for this variation.

Hair colour

Imagine two populations in which a high H^2 for hair colour is attained (Brauer and Chopra 1980). In one of these populations (population B) some individuals have a genetic predisposition to disproportionately seek out time spent in the ocean. Perhaps these individuals are excellent surfers due to genetic variants which aid them in strength and balance. This means that some individuals actively modify their environment so as to spend more time in the sun and saltwater—resulting in lighter hair. This activity contributes to the V_P of hair pigmentation in this population, and thus towards H^2: a case of active G–E covariance. This is presented in Fig. 3a. This may seem far-fetched, however, heritability has been estimated for recreational interests such as 'wilderness activities', 'physical fitness' (Lykken et al. 1993), and 'hunting and outdoor activities' (Hur et al. 1996).

H^2 in the population with this form of active G–E covariance is likely to be rejected by most interpreters, as time spent in the sun and saltwater appears to confound the results. Instead V_P would intuitively be attributed to differences in environmental factors. Alternatively, a population in which no G–E covariance is present (population A) represents a situation that would generally be considered a good study of the heritability of hair colour—where variation in hair pigmentation is caused by V_G without any covarying environmental influences. I will elaborate on why this is the case in "A gradation of intuitions" section.

Obesity

Obesity is another phenotype that has been empirically shown to be highly heritable (Stunkard et al. 1986), although the causal dynamics involved in the development of obesity and body mass differences are still largely unknown. One theory is that genetic differences affect 'nutrient partitioning' where some individuals lay down extra energy as fat or tissue due to genetic differences (O'Rahilly and Farooqi 2008). Other explanations involve a genetic influence on the regulation of appetite and energy balance, for example via the production of leptin—a hormone involved in relaying the fat status in the body to the brain (O'Malley and Stotz 2011). Additionally, psychological influences have been suggested as a causal factor. Depression is correlated with obesity, but it is unclear whether this is a cause or an effect of the condition (Stunkard et al. 2003). In line with the example given at the beginning of this article, a genetic basis of food preferences has also been suggested, as genetic associations with a preference for fat consumption, and resulting obesity, have been found in twin and family studies (Roberts et al. 2000).

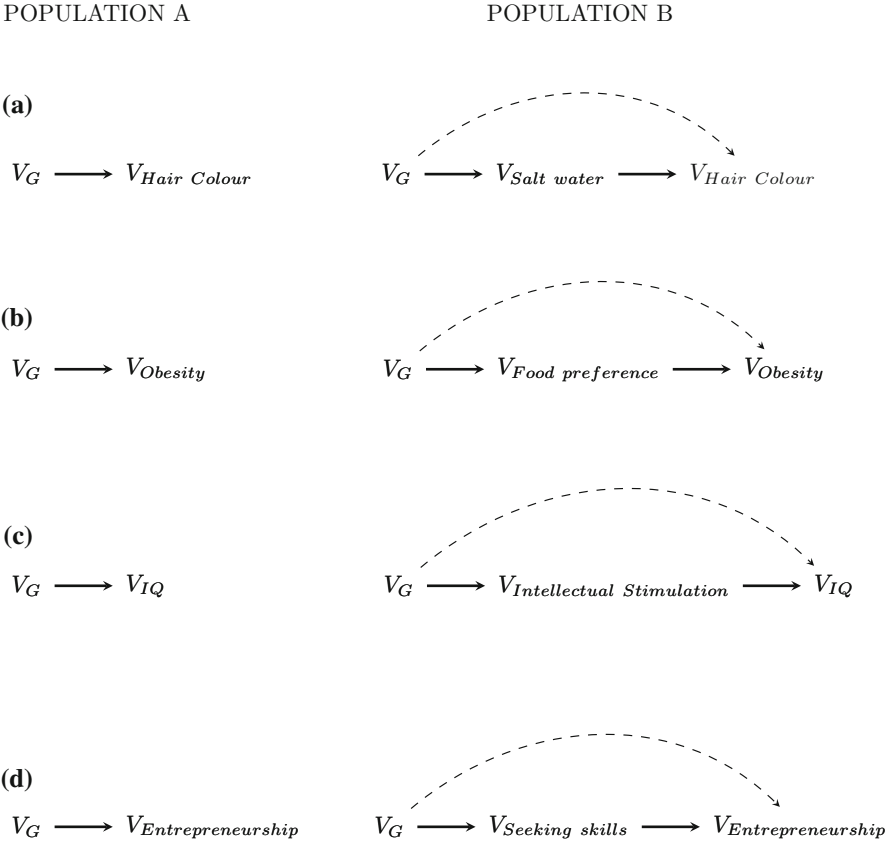

Fig. 3 In population A no G–E covariance is present, population B includes active G–E covariance. Intuitions about how to interpret the heritability estimates in Population B differ between different phenotypes (**a**–**d**)

To return to our two hypothetical populations, we can imagine that the heritability of obesity in population A is due to the 'direct effects' of nutrient partitioning—where heavier individuals lay down more energy as fat than less obese individuals. In population B, individuals are obese because they actively sought out environments that would contribute to weight gain. Perhaps they had a preference for fattier foods, and so surrounded themselves with a gustatory environment enabling increased fat consumption. Or perhaps they were genetically predisposed to suffer from depression, leading to over-eating. If the inclination to behave in any of these ways is at least partially genetically caused then active G–E covariance occurs: variation in genotype causes individuals to develop in a particular environment, and differences in that environment contributes to the variation in the measured phenotypic effect (Fig. 3b).

It seems intuitive to say that obesity is heritable in population A, in which no G–E covariance is present, but the interpretation of population B is less straightforward. In this population obesity may be instead attributed to environmental

variance, as it is the more direct cause of phenotypic variation, and as such the H^2 would be rejected.[9] How to interpret the results in population B seems to hinge on how one conceptualises obesity. Is it a purely physiological trait, or are causes of environmental modification acceptably included within the obesity concept? I shall return to this issue in "A gradation of intuitions" section.

Intelligence

Intelligence is the most widely studied phenotype in behavioural genetic research (Plomin and Spinath 2004), with a host of reviews estimating significant heritability estimates (Devlin et al. 1997; Plomin et al. 2008; Plomin and Spinath 2004).

"Active gene–environment covariance" section already demonstrated how intuitions about the heritability of intelligence vary amongst scholars when active G–E covariance is involved. To make this doubly clear, let us return to our two populations. In population A individuals with particular genes have higher IQ scores than others, as their genes affect their intelligence 'directly', perhaps through protein expression that influences neural development. This coheres with Block (1995) and Block and Dworkin's (1976) 'direct' heritability concept, and according to Jencks (1980, 730), is the assumed or received interpretation of the heritability of intelligence. Now consider population B, where individuals who score highly on the IQ test had a genetic propensity to modify their environments towards intellectual stimulation. They sought out books, took extra classes, and worked on problems which resulted in better literacy skills. In this population the high heritability estimate is (at least partially) a result of active G–E covariance. Variation in genotypes causes systematic variation in environments between the two genotype groups, producing variation between the groups for a phenotype (IQ). This is illustrated in Fig. 3c.

When considering 'common-sense' causal attributions, there is no concern about the heritability of IQ in population A. In population B, however, there is contention. "Active gene–environment covariance" section illustrated the division in interpretations for these kinds of cases. Some believe that cases like this should be interpreted in the same way as population A—where the resulting V_P is thought to be due to V_G. Others believe that the covariance should be partitioned separately (as Cov_{GE}), lowering the H^2.

Entrepreneurship

Lastly, let us consider the phenotype 'entrepreneurship'. Although this may seem an implausible candidate for a phenotype, let us for now assume the accuracy of some quantifiable entrepreneurial scale, and imagine a study in which a high H^2 estimate is reached (Nicolaou et al. 2008). In one population (A) high scorers on the entrepreneurial scale developed the phenotype in a 'direct' way. Their genes led to entrepreneurial behaviours and attributes in a biochemically-closed, neurologically-

[9] Interpretations of this example may also be biased due to prejudices against obese individuals (K. Stotz, personal communication).

mediated way. This is contrasted with population B, in which individuals who scored highly on the entrepreneurial scale were (at least partially) genetically influenced to actively seek out particular environments. These environments helped them to acquire the knowledge and skills that would aid in their entrepreneurial endeavours. Nicolaou and Shane (2009) have suggested that an active G–E covariance component is likely to explain their H^2 estimates, for instance through a genetic influence on educational and occupational preferences.

While there was an intuitive discord between the situations in populations A and B for the other three examples, the entrepreneurial case differs in that the situation in both populations seem to be acceptably heritable. That is, the H^2's attained in populations A and B both seem to accord with 'common sense' causal attributions. High scorers who actively sought out skills and knowledge is compatible with V_P being caused by V_G, resulting in a high H^2. This means the active G–E covariance cases are less problematic for this phenotype. This is because seeking out the skills and knowledge needed to be a successful entrepreneur seem to be part of *what it is to be* an entrepreneur. Thus the extended phenotype of skill seeking as part of the causal pathway to entrepreneurship appears acceptably encompassed by H^2. So when thinking about the heritability of entrepreneurship, implicit in the assessment of genetic and environmental causation is the assumption that individuals who display entrepreneurial traits would have reached them through some sort of active G–E covariance processes.

A gradation of intuitions

By mapping out the causal pathways that occur in population B for the hair colour, obesity, intelligence, and entrepreneurship examples one can see that they are causally congruous (Fig. 3), yet 'common sense' causal attributions differ. A high H^2 for hair colour clashes with common sense causal intuitions when active G–E covariance is present. In this example it is more intuitive to attribute the variation in hair pigmentation to an environmental factor: sun and/or saltwater exposure. Active G–E covariance cases of obesity also appear to be problematic for a heritability claim, although perhaps to a lesser degree. When individuals actively sought out particular environmental influences leading to weight gain (population B), the corresponding H^2 appeared more contentious than when heritable obesity differences were not due to environmental modifications (population A). Intuitions about intelligence are even less clear. This is evident from the existing debate surrounding active G–E covariance, where philosophers and biologists have long disagreed about how to causally attribute active G–E covariance. At the other end of the spectrum, active G–E covariance is acceptably included under V_G for the study of entrepreneurship, as in this case the covarying environmental modification appears to be *a part of* the phenotype under investigation.

These four phenotypes can be ordered along a continuum, with 'hair colour' at one extreme—where active G–E covariance cases (those in population B) are obviously problematic for the estimation of H^2, and 'entrepreneurship' at the other, where active G–E covariance cases are unproblematic, and elicit a response no different to non G–E covariance situations (those in population A). Thus for these

examples, the more 'acceptable' active G–E covariance cases are represented at the bottom of Fig. 3 and the less acceptable cases at the top.

So what accounts for the difference between these examples? I have already demonstrated that the effect is not due to any intrinsic causal characteristics of the system. One feature has been hinted at in the literature: phenotypes in which the co-varying environmental causes are generally considered as part of the 'self-realization' of the phenotype are those which are the least problematic for heritability analyses (Eaves et al. 1977; Jensen 1969). This accords with the four examples presented above. As hair colour is thought of as a physiological phenotype, sun-seeking is not a relevant or normal aspect of its realization. The 'normal' development of hair colour does not include excess time spent in the sun and saltwater, leading to the bleaching of hair follicles. However, seeking out resources and opportunities is highly relevant to entrepreneurship. These activities can be thought of as part of the normal expression or development of the trait itself, or, if genetically caused, as embodying its 'normal genetic development'. One could term them a 'natural manifestation' of the genotype for entrepreneurship, and thus are unproblematic as an expression of V_G.

Norms have already been shown to influence causal reasoning by playing a contrastive role (Knobe 2006). It follows from this that disagreement over causal attributions may occur when there are disagreements about the relevant norms that should be used as contrasts. For phenotypes like intelligence and obesity, the causal attributions for active G–E covariance cases are debated or unclear. This may be because the 'normal development' for these phenotypes is also unclear, debated, or underspecified. What is the normal development of obesity, or of intelligence? What environmental modifications are relevant to, or part of, these traits? Jinks and Fulker (1970, 323) make this point in reference to intelligence:

To what extent could we ever get a dull person to select for himself an intellectually stimulating environment to the same extent as a bright person might?

Because these questions are hard to answer without some controversy or debate, an interpretation of active G–E covariance is similarly difficult.

Thus to understand how active G–E covariance is interpreted in the heritability framework, one must look at the phenotype under study and ask questions about how this concept is conceived, and what kind of environmental interactions are expected or acceptable for its normal development. This changes the approach of handling cases of active G–E covariance from a general dichotomous one concerning what to do with cases generally, to one which may be assessed on a case-by-case basis.

The most important factor to consider for trait development in these cases is the 'seeking out' of stimulation or other environmental features. Environmental covariance appears to be acceptable when seeking out stimulation or environmental modification is relevant to or part of a phenotype's normal development. Another way to think of this is that environmental modification is relevant to a phenotype's realization when there is some active, *motivational* component to the concept of the phenotype itself. If there is some motivational component to the phenotype under study, then the activity of modifying one's environment can be thought of as a relevant expression of that phenotype. When this is the case, people should be happy

to include active G–E covariance as part of genetic variance. Conversely, if phenotypes do not include some sort of motivational feature, they should be less inclined to include active G–E covariance in a H^2 estimate, and instead want to partition this influence as a separate source of variance in the heritability model (Cov_{GE}).

This is evident across all of the above examples. Entrepreneurship is a phenotype in which the motivation to modify one's environment is paramount. Part of *what it is* to be an entrepreneur is to have the motivation to seek out resources and modify one's environment in a way that aids in economic success. On the other hand, physiological phenotypes like hair colour do not have any motivational component as part of their concept. So when motivational behaviour impacts upon the expression of the trait—even when these motivations are genetically caused, we tend to regard the resulting environmental variation as a separate source of variance. Obesity is a slightly trickier case. I believe that this sense of ambiguity is due to an inherent ambiguity in the obesity concept (see O'Malley and Stotz 2011). Those who believe that the obesity concept is purely a physiological one may not accept a H^2 estimate from an active G–E covariance case. However, those who include psychological, motivational and behavioural factors like appetite and food preferences as part of what it is to be obese or what the normal development of obesity entails may accept active G–E covariance as representative of genetic variance.

Recall that some regard the seeking out of intellectual stimulation as a natural component of the intelligence phenotype (Eaves et al. 1977). Under this view, the seeking out of stimulating environments which contribute to the development of the phenotype becomes part of the phenotype itself, and cases of active G–E covariance are permissible as heritable to the same extent as non-G–E covariance cases. Under this account the resulting phenotypic variance in both populations A and B would be included in H^2. However, those who think that intelligence is unrelated to the motivation to learn would maintain that any active G–E covariance skews the heritability estimate, and that the resulting phenotypic variance should be considered as caused by something other than V_G. This is the view of Block (1995), Block and Dworkin (1976), Gibbard (2001), Plomin et al. (1977), and Sober (2001). Therefore, disagreement about the intelligence case reduces to whether people think that part of what is being measured when heritability is estimated for intelligence is an interest in learning and practicing cognitive capacities, or simply an ability to perform the tasks.

As such, when attempting to interpret heritability estimates which include active G–E covariance, one must consider conceptual aspects of the phenotypic effect being measured. I contend that philosophical disagreement about causal attributions in active G–E covariance cases are in essence disagreements regarding how a phenotype should be defined. This moves the debate from one which concerns causal attributions and appropriate heritability models to one concerning the conceptual definition and 'normal development' of often ambiguous phenotypes like intelligence.

Conclusion

Debate about how to interpret heritability analyses involving G–E covariance is currently at a stand-still, with disagreement boiling down to differences in interpreters' 'common sense' causal intuitions. There remain two interpretive quandaries to be accounted for: (1) Why do interpretations differ between different types of G–E covariance (active, passive, reactive)? (2) Why do interpretations differ for some active G–E covariance cases? In regards to the first question I have demonstrated how factors like agency and blame, which have already been shown to influence causal reasoning in other contexts, account for the difference in how active and reactive cases are interpreted. The second question can be answered by considering the phenotype under investigation in each active G–E covariance example. Active G–E covariance may or may not be intuitively considered as part of a 'commonsensical' heritability estimate based on the kind of phenotype which is being measured. If the phenotype is one which has a large motivational dimension, or one in which environmental modification is regarded as part of its normal development, then active G–E covariance will not be problematic. This seems to be the case for some cognitive and behavioural phenotypes such as entrepreneurship. If the phenotype has a low motivational component, and/or environmental modification is not part of what is considered to be its 'normal development', then active G–E covariance cases appear problematic. This is the case for strictly physiological phenotypes, such as hair colour. Combined, a consideration of causal structures (distinguishing passive from reactive and active cases), notions of agency and blame (distinguishing active from reactive cases), and an examination of the phenotype under study can be used to account for differences in causal interpretations of G–E covariance.

Given that the interpretation of active G–E covariance cases are dependent upon the kind of phenotype studied, it is not surprising that phenotypes for which we have controversial, vague, or inexact definitions are the most problematic. Intelligence has been the case study for the majority of discussion of active G–E covariance, and intuitions elicited from this case study have differed between theorists. The interpretive intuitions evoked in this case have then been applied more generally by philosophers and biologists when trying to devise a rule to interpret active G–E covariance cases, leading to a more general disagreement about the interpretation of active G–E covariance. This may account for why the debate surrounding active G–E covariance is still unresolved. So far the debate about what to do with active G–E covariance has largely focused on behavioural and cognitive traits, particularly intelligence. I believe that this focus has obscured some of the key features of active G–E covariance, and has left the debate at a standstill. By considering other phenotypes, and identifying the role played by phenotypic concepts and ideas about normal development the debate can now be shifted to one concerning the specification of phenotypes.

This brings me to a third quandary regarding G–E covariance interpretations. How *should* these cases be regarded by behaviour geneticists? A full account of this issue is beyond the scope of this paper, but the analysis presented here provide

options for some initial first steps. Showing that both reactive and active cases of G–E covariance have identical causal structures will lead some to conclude that they should always be interpreted in the same way. This will be the case for those who believe that interpretation of behaviour genetic results should be based only on objective ontological considerations about the causal structure of the world. Whether or not to attribute the resulting variance to V_G, V_E, or COV_{GE} in these cases will require further argument. Others may be more motivated by concerns about commonsensicality in their interpretation, and factor in the phenotypic considerations from "Factors in causal reasoning: agency and blame" section, interpreting active cases using a semantic framework. For example, when dealing with situations where active G–E covariance is present, one must first decide how to define whichever phenotype is under study, and whether the concepts about that phenotype and its development include a motivational component, which may be expressed in the active alteration of the environment. This is still likely to spark debate as disagreements turn to how each phenotype is best defined and thought to normally develop. I do not provide an account here on how this might be resolved, it requires future philosophical work. But by beginning with this initial step the interpretative debate surrounding G–E covariance has advanced, moving from blind disagreement over whose causal intuitions match some 'true cause', to conceptual considerations of normal development and characteristic phenotypes.

Acknowledgments I would like to thank Karola Stotz, Paul Griffiths, Pierrick Bourrat, Chris Lean, and Richard Menary for helpful discussion and comments on earlier drafts of this work.

Appendix: Indirect causes

The significance of a blameworthy agent differentiating reactive and active G–E covariance cases relies on the premise that both have identical underlying causal structures, where genetic differences indirectly cause phenotypic differences via the environment. There are two possible concerns with this claim: (1) that reactive and active cases are not causally identical, (2) that one or both are not indirect causes. I shall address both of those concerns here.

Both types are represented in Fig. 2, where $V_G \rightarrow V_{P1} \rightarrow V_E \rightarrow V_{P2}$. This is just one way of representing these examples, as there are alternative ways of representing any causal system depending on how coarse or fine-grained the causal variables are partitioned. For instance, a representation of the reactive case in Fig. 2 could have included an additional variable, so that V_G causes variation in hair colour (V_{P1}), causing variation in the way that others behave towards these individuals (V_X), causing variation in environmental experience of those individuals (V_E), leading to variation in measured intelligence (V_{P2}). Some may be concerned that when the causal representations of these examples are made more specific the similarity in causal structure between active and reactive types disappears. Thus alternative ways of representing one type (active or reactive) of G–E covariance must be shown to apply to the other. Additionally, the indirect causal relationship should be preserved for both no matter the way in which variables are described. I

maintain that both types of G–E covariance are causally congruous, no matter the coarseness of grain of which they are represented. Note that the *nature* of some variables will be different between the active and reactive cases—this is emphasised in "Factors in causal reasoning: agency and blame" section. It is important to distinguish that differences in variable type is not the same as differences in causal structure.

To begin, I define a cause under the interventionist account of causation (Woodward 2003), where C causes E when an intervention on C, changing its value (e.g. from c_1 to c_2) results in a change of value in E (e.g. from e_1 to e_2). An intervention under this account refers to an idealised manipulation, where the change in value occurs without a change to the value of any other variable. This means that the change in the value of E is fully accounted for by the change in the value of C. For an indirect causal relationship, the transitivity of causation entails then when C causes I, and I causes E, then C causes E. This can be imagined as the value of C being set from c_1 to c_2, resulting in a change in I from i_1 to i_2, and a subsequent change in E from e_1 to e_2. According to Pearl (2001), an indirect effect (E) of a prior variable C is defined as an expected change in E that would occur when the value of C is held constant ($C = c_1$), while the intermediate variable (I) is changed to whatever value would have occurred given that C had been set to c_2, in this case $I = i_2$. Thus to determine if a variable E is indirectly caused by C, such that $C \rightarrow I \rightarrow E$, we must first look at what value the intermediate variable "I" would have been if C was intervened on, changing its value from c_1 to c_2.

I will begin by examining reactive cases. These examples are particularly controversial as it is an uncomfortable position to attribute differences in discrimination experienced (V_E) as caused by phenotypic differences of those experiencing the discrimination (V_{P1}). This is akin to stating that differences in skin colour are causes of racism within a population, which might lead some to deny that these examples represent indirect causation. Because of this I shall detail the logic behind this part of causal structure displayed in Fig. 2, where variation in hair colour (V_{P1}) causes variation in environmental experience (V_E), via variation in discriminatory behaviour by others, which I shall call V_X. Note that V_X is not displayed in Fig. 2, but can be thought of as a more fine-grained way of partitioning V_E.[10] This means an intervention on V_{P1}, setting it from $v_{p1(1)}$: no variation in hair colour phenotypes to $v_{p1(2)}$: some colour differences, would result in a corresponding change to the value of V_X (from v_{X1}: no variation in discriminatory behaviour by others, to v_{X2}: some variation in discriminatory behaviour from others). Note that in accordance with Woodward and Pearl's accounts, an intervention is made on V_{P1} and V_{P1} only, while other background variables such as societal beliefs and conventions are held constant. Under this account varying the hair colour variable produces a change in discriminatory behaviour towards this population, and so must be thought of as a cause of that behaviour. This is not to diminish the causal responsibility of other factors like societal mores in causing discriminatory behaviour, but shows that variation in hair colour is in this case *a* cause of these outcomes.

[10] It is assumed that the other links in the causal chain in Fig. 2 are generally accepted.

To show that V_{P1} is an indirect cause of V_E, we must now set V_X to the value that would have occurred if V_{P1} was set to $V_{p1(2)}$, that is, V_{X2}. If this is the case, and there is some variation in discriminatory behaviour by others, and the value of V_E would change from V_{E1}—no differences in environments experienced, to V_{E2}—in which environmental experiences vary according to the prejudice they have experienced. So I have shown that from a strictly causal point of view, variation in the hair colour phenotype causes variation in environments experienced, via variation in discriminatory behaviour from others. A possible reason for the discomfort in making this claim is a conflation of causal and normative responsibility. While variation in environmental experience is causally due to variation in phenotypes (and indirectly by variation in genotypes) within this population, this in no way entails moral culpability, nor connotations of immutability linking V_G, V_{P1} and V_{P2}. It remains the case that interventions on V_X or root causes of V_X in situations such as these are the most viable and ethically permissible strategies for eliminating discriminatory differences. Another reason for this intuition may be that variables like "discriminatory behaviour" involve other morally culpable agents, and the absence or presence of these kinds of agents can influence the way in which people causally reason, as demonstrated in "Factors in causal reasoning: agency and blame" section.

Turning to the active cases, these examples also fit the indirect causal structure described with the same number of variables: $V_G \rightarrow V_{P1} \rightarrow V_X \rightarrow V_E \rightarrow V_{P2}$. In the example from "Active gene–environment covariance" section genetic differences cause small initial differences in intelligence ($V_G \rightarrow V_{P1}$). These differences cause brighter children to stimulate themselves by seeking out environmental factors such as additional classes or problem sets. This makes V_X in this example something like "differences in attendance of extra classes". This is a different *type* of causal variable to "discriminatory behaviour from others", but that is not the point. In order to have causal similarity between active and reactive types of G–E covariance all that is necessary is that V_{P1} causes V_X. That is, if V_{P1} had a different value, e.g. if there was no variation in intelligence in the population, then given the same background conditions V_X would also change in value—and there would be no corresponding variation in class attendance. The next two steps, that differences in class attendance cause different environmental experience ($V_X \rightarrow V_E$) and that experience causes differences in measured intelligence ($V_E \rightarrow V_{P2}$) should be generally accepted.

So active and reactive cases of G–E covariance are causally identical because in both situations genetic differences cause phenotypic differences via variation in a second (often unrelated) phenotype, which, when all other variables are held fixed, in some way causes a differential environmental experience. How to appropriately describe the variables in a causal system will vary depending on each individual case, and may differ depending on one's intuitions about satisfactory causal explanations. The exact description of each case is not important to illustrate this point, as increasing or decreasing the number of variables in the chain leading from V_G to V_{P2} in reactive and active G–E covariance cases does not alter the fact that V_G indirectly causes V_{P2}, via V_E.

References

Alicke MD (1992) Culpable causation. J Pers Soc Psychol 63:368–378
Ambert AM (1997) Parents, children, and adolescents: interactive relationships and development in context. Haworth Press, New York
Baessler A, Hasinoff MJ, Fischer M et al (2005) Genetic linkage and association of the growth hormone secretagogue receptor (ghrelin receptor) gene in human obesity. Diabetes 54(1):259–267
Block N (1995) How heritability misleads about race. Cognition 56(2):99–128
Block N, Dworkin G (1976) IQ, heritability and inequality. In: Block N, Dworkin G (eds) The IQ controversy: critical readings. Pantheon, New York, pp 410–540
Brauer G, Chopra VP (1980) Estimating the heritability of hair colour and eye colour. J Hum Evol 9(8):625–630
Burger KS, Berner LA (2014) A functional neuroimaging review of obesity, appetitive hormones and ingestive behavior. Physiol Behav. doi:10.1016/j.physbeh.2014.04.025
Dawkins R (1982) The extended phenotype: the long reach of the gene. Oxford University Press, Oxford
Devlin B, Daniels M, Roeder K (1997) The heritability of IQ. Nature 388:468–471
Eaves LJ, Last K, Martin NG, Jinks JL (1977) A progressive approach to non-additivity and genotype–environmental covariance in the analysis of human differences. Brit J Math Stat Psychol 30(1):1–42
Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th edn. Longman, Harlow
Fisher RA (1918) The correlation between relatives on the supposition of mendelian inheritance. Philos Trans R Soc Edinb 52:399–433
Fuller JL (1979) Comment on Wahlsten. In: Royce JR, Moss LP (eds) Theoretical advances in behavioural genetics, Sijthoff & Noordhof, Alphen aan den Rijn
Gibbard A (2001) Genetic plans, genetic differences, and violence: some chief possibilities. In: Wasserman D, Wachbroit R (eds) Genetics and criminal behaviour. Cambridge University Press, Cambridge, pp 169–197
Hart HLA, Honore AM (1985) Causation in the law, 2nd edn. Clarendon Press, Oxford
Herrnstein RJ, Murray C (1994) The bell curve: intelligence and class structure in American life. The Free Press, New York
Hur YM, McGue M, Iacono WG (1996) Genetic and shared environmental influences on leisure-time interests in male adolescents. Pers Individ Differ 21(5):791–801
Jencks CS (1980) Heredity, environment and social policy reconsidered. Am Sociol Rev 45:723–736
Jencks CS, Smith M, Acland H et al (1972) Inequality: a reassessment of the effect of family and schooling in America. Basic Books, New York
Jensen AR (1969) How much can we boost IQ and scholastic achievement? Harvard Edu Rev 39:1–123
Jinks JL, Fulker DW (1970) Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behavior. Psychol Bull 73:311–349
Kitcher P (2001) Battling the undead: how (and how not) to resist genetic determinism. In: Singh RS, Krimbas K, Paul DB, Beatty J (eds) Thinking about evolution: historical, philosophical, and political perspectives. Cambridge University Press, Cambridge, pp 396–414
Knobe J (2006) Folk psychology, folk morality. Dissertation, Princeton University
Knobe J, Fraser B (2007) Causal judgement and moral judgment: two experiments. In: Sinnott-Armstrong W (ed) Moral psychology volume 2: the evolution of morality. MIT Press, Cambridge
Loehlin JC, DeFries JD (1987) Genotype-environment correlation and IQ. Behav Genet 17:263–277
Lykken DT, Bouchard TJ, McGue M, Tellegen A (1993) Heritability of interests: a twin study. J Appl Psychol 78(4):649–661
Lynch KE, Bourrat P Interpreting heritability causally. Philos Sci (accepted)
Maestripieri D, Mateo JM (2009) Maternal effects in mammals. The University of Chicago Press, Chicago
Nicolaou N, Shane S (2009) Can genetic factors influence the likelihood of engaging in entrepreneurial activity? J Bus Venturing 42(1):1–22
Nicolaou N, Shane S, Cherkas L, Hunkin J, Spector TD (2008) Is the tendency to engage in entrepreneurship genetic? Manag Sci 54(1):167–179
O'Malley MA, Stotz K (2011) Intervention, integration and translation in obesity research: genetic, developmental and metaorganismal approaches. Philos Ethics Humanit Med 6(2):1–14
O'Rahilly S, Farooqi IS (2008) Human obesity: a heritable neurobehavioral disorder that is highly sensitive to environmental conditions. Diabetes 57(11):2905–2910

Pearl J (2001) Direct and indirect effects. In: Proceedings of the seventeenth conference on uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., pp 411–420
Plomin R, Spinath FM (2004) Intelligence: genetics, genes, and genomics. J Pers Soc Psychol 86(1):112–129
Plomin R, DeFries JC, Loehlin JC (1977) Genotype–environment interaction and correlation in the analysis of human behavior. Psychol Bull 84(2):309–322
Plomin R, Loehlin JC, DeFries JC (1985) Genetic and environmental components of environmental influences. Dev Psychol 21(3):391–402
Plomin R, DeFries JC, McClearn GE, McGuffin P (2008) Behavioral genetics, 5th edn. Worth Publishers, New York
Roberts RC (1967) Some concepts and methods in quantitative genetics. In: Hirsch J (ed) Behavior genetics analysis. McGraw Hill, New York, pp 214–257
Roberts RE, Kaplan GA, Shema SJ, Strawbridge WJ (2000) Are the obese at greater risk for depression? Am J Epidemiol 152:163–170
Rutter M, Silberg J (2002) Environment interplay in relation to emotional and behavioral disturbance. Annu Rev Psychol 53(1):463–490
Sesardic N (2003) Heritability and indirect causation. Philos Sci 70(5):1002–1014
Sesardic N (2005) Making sense of heritability. Cambridge University Press, Cambridge
Sober E (2001) Separating nature and nurture. In: Wasserman D, Wachbroit R (eds) Genetics and criminal behavior. Cambridge University Press, Cambridge, pp 47–78
Stunkard AJ, Foch TT, Hrubec Z (1986) A twin study of human obesity. JAMA 256(1):51–54
Stunkard AJ, Faith MS, Allison KC (2003) Depression and obesity. Biol Psychiatry 54(3):330–337
Tabery J (2014) Beyond versus: the struggle to understand the interaction of nature and nurture. MIT Press, Cambridge
Wang X, Chen D, Niu T et al (2000) Genetic susceptibility to benzene and shortened gestation: evidence of gene–environment interaction. Am J Epidemiol 152(8):693–700
Woodward J (2003) Making things happen: a theory of causal explanation. Oxford University Press, Oxford