



Heritability and Etiology: Heritability estimates can provide causally relevant information

Jonathan Egeland

University of Agder, Norway

ARTICLE INFO

Keywords:

Heritability
Causation
Locality
Twin study
GWAS

ABSTRACT

Can heritability estimates provide causal information? This paper argues for an affirmative answer: since a non-nil heritability estimate satisfies certain characteristic properties of causation (i.e., association, manipulability, and counterfactual dependence), it increases the probability that the relation between genotypic variance and phenotypic variance is (at least partly) causal. Contrary to earlier proposals in the literature, the argument does not assume the correctness of any particular conception of the nature of causation, rather focusing on properties that are characteristic of causal relationships. The argument is defended against Lewontin's (1974) locality objection and Kaplan and Turkheimer's (2021) recent critique of Genome-Wide Association Studies (GWAS).

1. Introduction

The discipline of behavioral genetics aims to investigate and understand the relative influences of genes and environment on behavioral traits. With respect to such investigations, there are two fundamental questions that must be kept separate:

1. What proportion of an individual's phenotype P are genes responsible for, and what proportion is environment responsible for?
2. What proportion of the variance of phenotype P is genetic variance responsible for, and what proportion is environmental variance responsible for?¹

There is general agreement among researchers working on conceptual or methodological issues in behavioral genetics that the first question cannot be answered (e.g., Dowens & Lucas, 2020; Griffiths et al., 2005; Lewontin, 1974; Pearson, 2007; Sober, 2001). For example, if a person has a bodyweight of 70 kg, it does not make any sense to say that either genes or environment is responsible for a certain proportion of the person's bodyweight (such as that 50 kg are due to genes and 20 due to environment). Since both genes and environment necessarily contribute in intricate interaction with each other to the development of the phenotype in question, it is impossible to partition the phenotype into portions that are due only to genes or environment respectively. Genes and environment, nature and nurture, function as interwoven strands of

thread in the un-untieable Gordian knot that is the development of any phenotypic trait (cf. Bateson, 2001, p. 565; Nuffield, 2002, p. 40).

On the other hand, when it comes to the second question opinions differ as to whether it can be answered using current statistical methods and technologies. More specifically, the disagreement has to do with how heritability measures should be interpreted, and whether they can provide any information about the causal effects of genetic variance on phenotypic variance. This paper develops an argument to the effect that heritability measures indeed can provide causally relevant information about the sources of trait variance and, moreover, it shows that the argument is not vulnerable to either Lewontin's (1974) locality objection or Kaplan and Turkheimer's (2021) recent critique of Genome-Wide Association Studies (GWAS).

The paper is structured as follows. Section 2 lays the groundwork by explaining the basics of heritability estimation and why there is so much disagreement about how such estimates should be interpreted. Section 3 develops the argument that heritability measures can provide causally relevant information. Section 4 shows that although heritability measures are contextual and local, it does not follow that they are scientifically useless and without causal relevance. Section 5 shows that Kaplan and Turkheimer's recent critique of GWAS is unsound. Section 6 concludes and offers some reflections on the threats posed by gene-environment (G-E) interaction, G-E covariation, and indirect genetic effects (IGEs).

E-mail address: jonathan.egeland@uia.no.

¹ The first question is about the phenotype of a particular individual, whereas the second is about the occurrence of certain phenotype in a particular population (since variance is a population statistic). Cf. Tal's (2009) discussion of the first question and of an alternative version of the second question that also is focused on the phenotype of a particular individual.

<https://doi.org/10.1016/j.paid.2022.111896>

Received 21 July 2022; Received in revised form 30 August 2022; Accepted 1 September 2022

Available online 6 September 2022

0191-8869/© 2022 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

2. Heritability: analyzing trait variance

Heritability estimates are often calculated using a statistical method known as the analysis of variance (ANOVA). ANOVA is based on a linear model, which in the case of heritability estimation assumes that the variance of a phenotypic trait V_P is a linear function of genotypic variance V_G and environmental variance V_E (given that there is no G-E interaction or covariation):

$$V_P = V_G + V_E$$

And the most common way of defining a phenotypic trait's heritability (H^2) is as the ratio of V_G to V_P (Plomin, DeFries, McClearn, & McGuffin, 2008)²:

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

H^2 is a statistical measure of *broad sense heritability*, which is the estimated proportion of phenotypic variance that is due to genetic variance. However, there is also another heritability measure known as *narrow sense heritability* (h^2), which is the estimated proportion of phenotypic variance that is due to additive genetic variance. Additive genetic variance is simply the proportion of phenotypic variance that is due to the additive effects of genes. The other sources of genetic variance are dominance variance, epistatic variance, and variance due to assortative mating. For the purposes of this paper, we will focus on heritability in the broad sense.

From what has been said above, it may seem intuitive that non-nil heritability measures provide information about genetic causation. After all, discovering that a certain trait has a H^2 of, say, 0.5 (which is not uncommon for behavioral traits, Plomin, DeFries, Knopik, & Neiderhiser, 2016; Polderman et al., 2015) means that genetic variation (measured in phenotypic units) explains 50% of trait variation in the population that is being studied. And it may not seem a stretch to think that some of the genetic variation is *causally* responsible for some of the phenotypic variation (e.g., Sesardic, 2005, p. 82ff.). However, there are many objections to this idea. One is that the definition of H^2 given above only says that there is an associative relation between the terms V_G and V_P —and as we have all been taught in statistics 101, association does not imply causation (Turkheimer, 2016). Another objection is that the definition of V_P given above is incomplete. More specifically, it has been noted that one cannot simply assume that there is not any G-E interaction $V_{G \times E}$, or any G-E covariation $2COV(G, E)$.³ The definition of phenotypic variance should therefore be amended as follows⁴:

$$V_P = V_G + V_E + V_{G \times E} + 2COV(G, E)$$

And, moreover, many commentators have argued that covariation between genotypes and environments constitutes a challenge to the claim that heritability analyses of trait variance can provide causally relevant information (Block, 1995, pp. 118–121; Block & Dworkin, 1976a, p. 480; Feldman & Lewontin, 1975, p. 1164; Jencks, 1980, pp. 726–730; Kaplan & Turkheimer, 2021, p. 61; Sober, 2001, pp. 72–75).

A famous thought experiment from Jencks et al. (1972) illustrates

² For more on the various ways of defining heritability, see Bourrat (2015); Dowens and Lucas (2020); Falconer and Mackey (1996); Godfrey-Smith (2009); Hartl and Clark (1997); Jacquard (1983); Tal (2009).

³ G-E interaction occurs when different genotypes don't respond to the environment in the same manner, or when the environment has differential effects that contingently depend on individuals' genotypes; G-E covariation occurs when certain environmental factors are associated with genetic propensities in a population.

⁴ The formula is still somewhat simplified: V_E is usually divided into between-family variance and within-family variance, and the formula should also include an error term representing variance due to measurement error. However, for present purposes, it is only necessary to keep in mind the simplified version of the formula.

why large G-E covariation renders any causal interpretation of heritability unjustified. Imagine that there is a population in which children with red hair experience systematic discrimination, and they are denied access to education. In this population, red haired children will on average perform worse on measures of intellectual aptitude and achievement. Moreover, since there is large covariation between genes associated with red hair and experiencing discrimination, the genetic variants associated with red hair will be predictive of low scores on measures of intellectual aptitude and achievement. However, according to some of the aforementioned commentators, heritability measures of aptitude and achievement cannot be trusted to be indicative of causal relationships in this population, since on any intuitive understanding of what it means for a gene to have a causal effect on a trait, one cannot say that the genes associated with red hair are causally responsible for low scores. To the contrary, it should be obvious that the etiological root of the relatively low scores of red-haired children is the discrimination they experience—or at least so the argument goes.

As things currently stand, there appears to be a lot of disagreement concerning the interpretation of heritability measures and whether they can provide causally relevant information. The next section will bracket issues having to do with G-E covariation and develop an argument for the claim that heritability measures are not causally irrelevant. After that, the argument is defended against a couple of prominent objections from the literature.

3. The argument for causal relevancy

It is not uncommon for both proponents and opponents of the claim that heritability can be given a causal interpretation to rely on particular views about the nature of causation (cf. Oftedal, 2005). For example, Lewontin (1974) and several of his followers (e.g., Block, 1995, p. 24; Block & Dworkin, 1976a, p. 482; Kaplan & Turkheimer, 2021; Keller, 2013) argue that ANOVA is not a useful method for discovering causal relationships since ANOVA only provides associative information, and knowledge about genetic causation requires knowledge of the exact “process”, “function”, or “mechanism” by which genes produce their phenotypic effects. Indeed, as Lewontin has put it, knowledge of genetic causation requires that one can “provide a detailed molecular analysis of the chain of causation between nucleotide substitution and cell development and function” (Lewontin, 2006, p. 537).

Moreover, the same tendency is also found among those who support causal interpretations of heritability. But whereas the opponents rely on so-called process conceptions, the proponents tend to understand causation in terms of probability or difference-making. Examples of this can be found in recent work by Bourrat (2019, 2020); Lynch and Bourrat (2017); and Tal (2009), where it is argued that heritability estimates can be given a causal interpretation since they show that certain genes increase the probability of having a phenotypic trait with a value that deviates from the mean value of said trait.

However, there are problems with both strategies, stemming from the fact that they (in part) wed their claims about whether heritability measures can be interpreted causally to particular conceptions about the nature of causation. Two problems, more specifically, are that both kinds of conception face intuitive counterexamples, indicating that causation cannot be analyzed in terms of “process” or “probability” alone, and that there is not at the moment a consensus about what the correct theory of causation looks like.⁵ (For more on the different theories of causation, as well as their strengths and weaknesses, see the review by Schaffer, 2016.) Continuing, the argument of this paper will not assume the correctness of any particular conception of the nature of causation. Rather, it will focus on properties that are *characteristic* of

⁵ Another problem that primarily threatens the views of the opponents, is that they seldom spell out their assumptions about causation in any detail. This should be contrasted with the approach of (e.g.) Bourrat (2019, 2020), who explicitly relies on Woodward's (2003) interventionist account of causation.

causal relationships, without presupposing them to be either necessary or sufficient.

Briefly put, the argument is that since a non-nil heritability measure tells us that the relationship between V_G and V_P satisfies certain characteristic properties of causation, it increases the probability that the relation is (at least partly) causal—whatever the nature of causation really is. The properties that will be focused on are association, manipulability, and counterfactual dependence.⁶ Consider association first.

A couple of variables stand in an associative relation to each other just in case certain values of one variable make certain values of the other variable more likely. For example, since UV radiation exposure increases the likelihood that one will develop skin cancer, UV radiation exposure and skin cancer are associated with each other. Another property that often is relied upon to get a grip on causation is manipulability (Woodward, 2016). It is not uncommon to understand causation in terms of the idea that the manipulation of a cause (and no other variable)⁷ regularly will result in the manipulation of an effect. The manipulation need not be experimental or to have actually occurred; rather, it is enough that it can happen in principle. A third characteristic of causal relationships is counterfactual dependence, meaning that the cause is counterfactually necessary for the effect (Menzies & Beebe, 2020). More specifically, a variable Y counterfactually depends on a variable X just in case there are alternative, possible values for Y and X, such that if X were to have an alternative value, then so would Y.

Association, manipulability, and counterfactual dependence are characteristic properties of causal relations and, moreover, that is precisely why we acquire causal knowledge by relying on information concerning whether, and to what extent, certain relations of interest satisfy these properties. In fact, the evidential support of causal claims in the sciences is typically considered proportional to the degree to which the data indicate that the aforementioned properties are satisfied. Consider as an example the claim that smoking causes cancer. The consensus view is that smoking indeed causes various types of cancer, and that we *know* this to be true. But how did we acquire this knowledge? We know it because a number of well-designed studies have been published, where it has been found, based on analyses of large datasets, that smoking and cancer are significantly associated even after controlling for possible confounders.⁸ The associations are interpreted as evidencing causal relationships since the statistical analyses are based on large and representative samples from different populations living under varying environmental conditions, and they remove the effects of other relevant factors. This means that smoking is associated with cancer, and since changing the value of *only* the smoking variable would lead to a change in the value of the cancer variable, the relationship also has the properties of manipulability and counterfactual dependence. Although we may not know everything about the mechanisms by which the cancers are developed, or about what exactly the necessary and sufficient conditions for causation are, learning that a relationship between two variables satisfies the three aforementioned properties does increase the probability that it is causal—sometimes to such an extent that we can *know* that the relationship is one of cause and effect.

The situation is in many ways analogous when it comes to heritability. When $H^2 > 0$, there is an association between V_G and V_P . Moreover, although no ideal intervention is performed in either classical twin studies or GWAS, the relation between genetic variance and phenotypic variance does appear to satisfy manipulability and counterfactual dependence since V_P is defined as a linear function of V_G and V_E

—meaning that a change in the value of V_G only, should induce a change in the value of V_P . It is of course always theoretically possible that some confounding factor (such as population stratification or assortative mating) is responsible for the change in trait variation, or that there is a very large G-E interaction effect that obscures the relationship between V_G and V_P (cf. Lewontin, 1974, p. 406), but the likelihood of this possibility is greatly reduced when the samples on which the estimates are based are large, appropriate statistical techniques are used (for more on this, see Young, Benonisdottir, Przeworski, & Kong, 2019), and there is no evidence for large G-E interaction.

Now assuming that there is not any large G-E covariation or interaction,⁹ the same reasoning can be used to illustrate that non-nil heritability measures increase the probability that genotypic differences cause phenotypic differences. Since heritability measures satisfy certain characteristic properties of causation (i.e., association, manipulability, and counterfactual dependence), and cases where those conditions are satisfied constitute a proper subset of all possible cases that exist with respect to the relationship between certain genes and traits—one that includes all, or at least most, cases of genetic causation—knowledge that a certain phenotypic trait has a non-zero heritability value increases the probability that genes play a causal role in the development of individual differences in said phenotype. The reasoning is illustrated in Fig. 1.

Let's summarize. Since we do not have direct epistemic access to the causal structures of the world, we have to make inferences about causation based on whether, and to what extent, certain relations of interest satisfy properties that are characteristic (and thereby indicative) of causation. Three such properties are association, manipulability, and counterfactual dependence. Moreover, as has been demonstrated above, heritability measures typically do satisfy these properties, which means that they can provide causally relevant information. However, it does not follow that non-nil heritability estimates always justify causal inferences. Whether they do will have to be judged on a case-by-case basis,

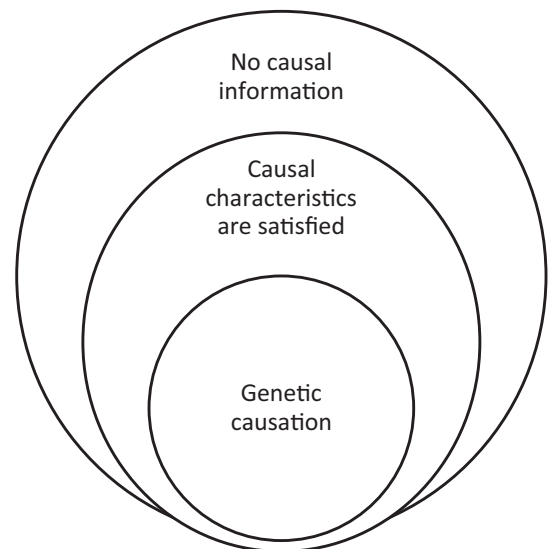


Fig. 1. The outer circle represents all possible ways in which genotypes and phenotypes can be related. The middle circle represents all the ways in which genotypes and phenotypes can be related when the properties of association, manipulability, and counterfactual dependence are satisfied. The inner circle represents all the ways in which variation in genotypes can cause variation in phenotypes.

⁶ An important implication is that Turkheimer (2016) and others are mistaken when they claim that heritability simply provides information about genotype-phenotype correlation.

⁷ A manipulation that only changes the value of one variable is sometimes called an ideal intervention (Woodward, 2003).

⁸ For a few examples, see Boffetta (2008); Lee, Forey, and Coombs (2012); Pesch et al. (2012); Sasco, Secretan, and Straif (2004).

⁹ Scenarios involving G-E covariation or interaction will be discussed in Section 6.

and in the light of evidence pertaining to the degree to which there is G-E covariation or interaction.

4. Lewontin's locality objection

Critics of the claim that heritability estimates can provide causal information usually rely on arguments presented in [Lewontin's seminal \(1974\)](#) article. One of Lewontin's most influential arguments is the so-called locality objection, which points out that H^2 is a population statistic and infers that it cannot be applied to any other population or any other measurement condition. This is how he puts it:

That is, the linear model is a *local analysis*. It gives a result that depends upon the actual distribution of genotypes and environments in the particular population sampled. Therefore, the result of the analysis has historical (i.e., spatiotemporal) limitation and is not in general a statement about *functional* relations ([Lewontin, 1974](#), p. 403).

Moreover, since Lewontin assumes that a causal analysis or explanation requires knowledge of the exact function or mechanism by which a cause produces its effects, as indicated by the last sentence above,¹⁰ it follows that ANOVA cannot provide any information about genetic causation.

The locality objection has been tremendously influential, and it has been reiterated by a number of scientists and philosophers who agree that because heritability is a population statistic it can be justifiably inferred that it cannot be applied to any other population or any other measurement condition (e.g., [Block & Dworkin, 1976b](#), pp. 486–487; [Daniels, Devlin, & Roeder, 1997](#), p. 54; [Nelkin & Andrews, 1996](#), p. 13; [Rutter, 1997](#), p. 391; [Rutter, 2002](#), p. 2; [West-Eberhard, 2003](#), pp. 102–103), and who also think that ANOVA cannot provide causally relevant information because it does not give any insight into the function or mechanism by which genes produce their phenotypic effects (e.g., [Block, 1995](#), pp. 117–119; [Block & Dworkin, 1976b](#), p. 482; [Kaplan & Turkheimer, 2021](#), p. 61ff.; [Keller, 2013](#)). However, there are five reasons why the locality objection does not have the dialectical force that many commentators have believed, and why it fails to threaten the argument from the previous section.

First, just because heritability is a population statistic, it does not follow that heritability estimates cannot provide any information about the relative contributions of genes and environment to individual trait differences in other populations than the one that has been sampled, or in other measurement circumstances. Determining the extent to which heritability estimates are generalizable is ultimately an empirical question, meaning that it cannot be answered by reflection or conceptual analysis alone ([Bouchard & Loehlin, 2001](#), p. 247).

Second, there is some empirical evidence supporting the generalizability of heritability. For example, high heritability values for general intelligence have been found in different countries, with their own particular cultures and environmental conditions, from different continents ([Knopik, Neiderhiser, DeFries, & Plomin, 2017](#), pp. 170–173). Moreover, it is to be expected that heritability measures from similar contexts and similar measurement conditions will not be altogether unlike each other. For a more detailed discussion of this issue, see [Sesardic \(2005, pp. 78–86\)](#).

Third, even if it were true that heritability measures never can provide any information or indication about how heritable a trait is in other populations than the one that has been sampled, or in other measurement circumstances, nothing follows concerning the issue of causal interpretation. Just because the ratio of V_G to V_P may be context-dependent, it can still be the case that it says something about the causal contribution of genetic variance to phenotypic variance *in the population from which the sample has been gathered* ([Tal, 2009](#), pp. 90–91).

In general: the nongeneralizability of an associative relationship does not imply that the relationship is not causal in the context where the association is present.

Fourth, it should be noted that Lewontin actually disagrees with this claim. He tells his readers that one should avoid “confusing the spatio-temporally *local* analysis of variance with the *global* analysis of causes” ([Lewontin, 1974](#), pp. 410, italics added), the latter of which requires knowledge about “functional relations” ([Lewontin, 1974](#), p. 403) that hold true of “the entire spectrum of causal relations” ([Lewontin, 1974](#), p. 407). However, this is a very radical claim—one that would (if it were true) undermine many, if not most, causal claims made in the sciences. For example, in scientific disciplines such as medicine, biology and psychology, we are interested in understanding how things work under relatively normal parametric conditions. This does not mean that such disciplines cannot discover causal relations, but rather that the causal relations that we know to hold true in most contexts may break down in more rarely occurring, or (e.g.) counterfactual, contexts. For example, just because we haven't investigated the functional relationship between sugar consumption and diabetes under “the entire spectrum” of environmental conditions, it would be irrational to claim that we cannot know that sugar consumption causes diabetes. A relatively local analysis of causes can be compatible with ignorance about global function (cf. [Haldane, 1938](#), p. 34).

Fifth, it is wrong to assume that knowledge about causation requires knowledge about “function” or “mechanism”. For example, we know that having a third copy of chromosome 21 causally contributes to lower IQ, even though there is a lot about function or mechanism that we don't know. Moreover, blaming heritability estimation for not providing insight into mechanism or function is like blaming the Beck Depression Inventory (used for the measurement of depression severity) for failing to say anything about *why* people become depressed. The point is simply that heritability measures can provide causal information without saying, or even purporting to say, anything about how genes influence phenotype development.

Taken together, these reasons demonstrate that Lewontin's locality objection does not undermine the argument of this paper.

5. The quincunx analogy

A GWAS is performed by searching for single-nucleotide polymorphisms (SNPs)—i.e., substitutions of single nucleotides—associated with a trait of interest. In cases of synonymous substitutions, the mutations do not lead to any phenotypic difference, but in other cases they do. The purpose of a GWAS is to identify SNPs that are associated with a certain trait by separating those who have the trait and those who do not have it into different groups, and by identifying variants that are more common in the group with the trait of interest. Now the SNPs that are identified are not necessarily associated with the trait themselves, as it is possible that they are located close to regions of the genome that are associated with the trait, and that usually are inherited collectively—in which case the variants in question are said to be in linkage disequilibrium. That said, there are ways of extenuating this problem, and weighted sums of SNPs that are found to be associated with the trait take the form of polygenic scores (PGSs) that function as predictors of the trait. Moreover, the variants on the basis of which a PGS is calculated are sometimes said to be causal (e.g., [Yang, Zeng, Goddard, Wray, & Visscher, 2017](#), p. 1305).

In a recent paper, [Kaplan and Turkheimer \(2021, p. 61ff.\)](#) have argued that GWAS face similar problems as those of the ANOVA approach to heritability estimation already pointed out by [Lewontin \(1974\)](#). They present an argument by analogy, focusing on Galton's quincunx (better known as the “Galton board”). The quincunx is a vertical board with rows of pins. When a ball is dropped into the top of the quincunx, it bounces either left or right when it hits the pins, eventually landing in one of the bins at the bottom of the machine (see [Galton, 1894](#), for a more detailed description). The pins, we are told, are

¹⁰ Cf. [Lewontin \(1974, p. 409; 2006, p. 537\)](#).

“difference makers”, in the sense that their relative placement is associated with certain outcomes (i.e., the distribution of balls in the bins at the bottom); but, for the purposes of their thought experiment, it is assumed that the walls of the quincunx are opaque, so that it is impossible to observe which path the ball travels.

Next, Kaplan and Turkheimer ask us to imagine a population of quincunxes with identifiable varieties of pin placements. If a particular pin placement at a particular location is more frequent among those quincunxes in which the ball landed to the left, then it is likely to have a left-bias: it will be predictive of left-biased outcomes. Moreover, the situation is analogous to GWAS in the following ways: (1) a particular variant of pin placement at a particular location in the quincunx functions in the same way as an individual SNP in the genome: just as pin placements are associated with certain ball distributions, so are SNPs associated with certain phenotypic traits; (2) the predictions made on the basis of the bias that exists in a population of quincunxes corresponds to PGSs: they are indicative of ball distributions and phenotypic outcomes respectively.

In their discussion of the analogy, Kaplan and Turkheimer make a number of points—most of which have been made by previous commentators (e.g., concerning G-E interaction, G-E covariation, and reasons why PGSs in one population may not be equally predictive in other populations) and are not directly related to the analogy. But the most important insight, we are told, is this: knowledge about how particular pin placements (or SNPs) are associated with certain ball distributions (or phenotypic traits), does not contribute to our “understanding where a particular ball ended up”. And the reason is that

Since a ball only interacts with a small minority of the pins in any particular trial, most of the time the pin in question will have been entirely irrelevant to the ball's path. Even when it was relevant, however, if all we know is that, at that location, there is a pin variant with tendency towards one direction, we can't know (except in cases of 100% bias) if the ball in fact took the more typical path, or if it took the path that was less likely (Kaplan & Turkheimer, 2021, p. 64).

And this is important since it supports the idea that GWAS cannot provide any relevant information about causation, or even about individual differences:

The associations discovered by GWAS (and related technologies) are unlikely to provide any meaningful basis for explaining variation in individuals; still less do they themselves reliably point towards causes of individual outcomes (Kaplan & Turkheimer, 2021, p. 60).

There are, however, two problems with this argument. The first problem is that it is not necessary that a particular pin/SNP causally contributes to the outcome in the case of a particular ball/person, or that we know with 100% certainty what the causal path taken is, in order for us to know that certain pin placements/SNPs satisfy important properties (i.e., association, manipulability, and counterfactual dependence) that increase the probability that they are causally related to certain outcomes. Precise knowledge about which causal path has been taken is not a necessary condition for gleaning information to the effect it is *some-what probable* that *some such* path has been taken. When we learn that certain SNP variants are associated with a certain trait, it is rational to somewhat increase our credence that they are causal—not to conclude that they cannot provide information relevant for understanding the outcomes observed since it is impossible to know exactly whether, or how, a particular SNP has contributed to a particular phenotype.

Now one may reasonably question whether Kaplan and Turkheimer really do think that being provided with causally relevant information requires knowledge about exactly what the causal “path” from SNP to phenotype looks like. Here are a couple of quotes showing that they do:

To understand the causal role that a gene plays in the development of a trait is therefore to understand when (under what conditions) and

how it is transcribed, and how the products are used across the development of the trait in question (Kaplan & Turkheimer, 2021, p. 61).

If we understand Lewontin's projects to be about understanding how individuals develop the traits that they have, and understanding why the distribution of traits in a particular population is the way that it actually is, the kinds of results given to us by GWAS/PGS will be of no more value to us than the kinds of ANOVA-based quantitative genetics research that he was criticizing. The analogy of the quincunx helps us to see why (Kaplan & Turkheimer, 2021, p. 68).

And this is the source of the second problem. Since Kaplan and Turkheimer indeed are followers of Lewontin's projects, they assume that a necessary condition for causal knowledge (and even just being provided with causally relevant information about genetic causation) is knowledge of the exact mechanism by which genes contribute to the development of a trait. However, as we have seen in the previous section, this is clearly setting the bar too high, and reflection on relevant examples illustrates why: Lacking an awareness of the mechanism by which caffeine functions as an adenosine antagonist, blocking the action of adenosine on its receptors, does not prevent a child from learning (either by testimony or experience) that caffeine consumption has the effect of reducing drowsiness. Inferring that it does from the Lewontonian position would appear to be a *reductio* against said position.

Moreover, insisting that knowledge of mechanism is necessary for knowledge of causation is really to endorse a sort of skepticism about the behavioral sciences, as it would threaten many, if not most, of their claims to causal knowledge. Or as Turkheimer, Goldsmith, and Gottesman (1995, p. 149) once rhetorically asked: “If knowledge of mechanism were required prior to investigation of relationships between predictor and outcome, how much of behavioral science would be disallowed?” However, this is clearly an unacceptable consequence. We know that trisomy 21 (i.e., down syndrome) causes lower IQ, even though we do not “understand when (under what conditions) and how [the relevant genes are] transcribed, and how the products are used across the development of the trait in question”.

6. Conclusion

This paper has argued that heritability measures can provide causally relevant information. Since a non-nil heritability measure tells us that the relationship between genetic variance and phenotypic variance satisfies certain characteristic properties of causation, it increases the probability that the relation is causal. Furthermore, the argument was defended against Lewontin's locality objection and Kaplan and Turkheimer's recent quincunx analogy.

However, critics of heritability estimation may very well claim that the most important objections—namely, G-E interaction and G-E covariation—haven't been addressed. This is true, and I want to make a few closing remarks with respect to these objections. First, when G-E interaction and G-E covariation are relatively small or moderate, they do not swamp out the main effects of genotypes and environments (Sesardic, 2005, ch. 2–3).¹¹ Second, the question of how additive the relations between genotypes and environments are for human traits is ultimately an empirical one, and only very few significant G-E interaction effects have been discovered and replicated (Gauderman et al., 2017; McGue &

¹¹ It should be noted that Lynch and Bourrat (2017) recently have argued that both active and reactive G-E covariance should be included in the V_G term. They argue that this is the only consistent way of interpreting heritability estimates in a causal manner.

Carey, 2017). Knopik et al. (2017, p. ch. 8) provide a useful summary of the literature, explaining that a large proportion of reported $G \times E$ effects do not replicate (cf. the literature review by Duncan & Keller, 2011), that most replicated effects have to do with non-cognitive traits,¹² and that genuine $G \times E$ effects usually are small enough that they do not obscure the main effects of genotypes and environment.¹³ Third, even if it turns out that nonadditivity is the rule rather than the exception,¹⁴ it does not undermine the argument of this paper. Since the argument only claims that heritability measures *can* provide causally relevant information—not that they provide causally relevant information under all (or even most) measurement conditions, or that they always justify causal inferences—they are only likely to do so in cases where we do not know that G-E interaction or G-E covariation does not leave room for readily interpretable main effects.

Lastly, it is worth mentioning that recent work in sociogenomics evidencing IGEs—i.e., effects whereby phenotype expression is influenced by the genotypes of other conspecifics—may complicate the issue of heritability estimation and causal inference. Indeed, observable effects of genetic nurture (Kong et al., 2018) and social epistasis (Domingue et al., 2018) may increase H^2 values for certain traits, even though this is not solely due to the individuals' own genotypes. Some may argue that this weakens the plausibility of genetic inferences, since genetic effects must be endogenous. However, a problem with this position is that an important lesson of the gene-centered view of evolution is that our common-sense conceptual distinctions between the individual on the one hand, and the social on the other, may not be entirely adequate for making sense of biological reality. Just as the genotype of an individual can have effects on *extended phenotypes*, there does not appear to be any scientific reason as to why it should not be possible for the phenotype of an individual to be influenced by its *extended genotype*, or why this should not count as genuine genetic causation. The organismal world may not always “respect” our intuitive conceptual distinctions, developed for dealing with non-scientific, everyday matters, but, as the history of science teaches us, our conceptual framework and the types of sense-making that it enables can be reformed in order to improve our understanding.

CRedit authorship contribution statement

Jonathan Egeland is the sole author of this paper.

For helpful comments, he thanks Pierrick Bourrat and Neven Sesardić.

Declaration of competing interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Data availability

No data was used for the research described in the article.

¹² An interesting exception is the Scarr-Rowe effect, whereby the heritability of general cognitive ability depends on socio-economic status. The effect replicates in the USA (even using a PGS for educational attainment, Peñaherrera-Aguirre, Woodley, Sarraf, & Beaver, 2022) but not in other Western nations (Tucker-Drob & Bates, 2016), or in more environmentally deprived areas of the world (Hur & Bates, 2019).

¹³ There does however appear to be significant, non-zero G-E covariation, which implies that some of the genotype-phenotype association is not due to direct genetic effects (e.g., Kendler & Baker, 2007; Krapohl & Plomin, 2016; Okbay et al., 2022; Plomin, 2014).

¹⁴ It is should be noted that although individual $G \times E$ effects are small, it is possible that they (just like individual SNP effects, see Chabris, Lee, Cesarini, Benjamin, & Laibson, 2015) have larger influences on trait variance when considered collectively, as pointed out by McGue and Carey (2017).

References

- Bateson, P. (2001). Where does our behaviour come from? *Journal of Biosciences*, 26(5), 561–570. <https://doi.org/10.1007/BF02704754>
- Block, N. J. (1995). How heritability misleads about race. *Cognition*, 56(2), 99–128. [https://doi.org/10.1016/0010-0277\(95\)00678-R](https://doi.org/10.1016/0010-0277(95)00678-R)
- Block, N. J., & Dworkin, G. (1976a). *The IQ controversy: Critical readings*. Oxford, England: Pantheon.
- Block, N. J., & Dworkin, G. (1976b). IQ, heritability and inequality. In N. J. B. G. Dworkin (Ed.), *The IQ controversy: Critical readings*. New York: Pantheon.
- Boffetta, P. (2008). Tobacco smoking and risk of bladder cancer. *Scandinavian Journal of Urology and Nephrology*, 42(sup218), 45–54. <https://doi.org/10.1080/03008880802283664>
- Bouchard, T. J., & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics*, 31(3), 243–273. <https://doi.org/10.1023/A:1012294324713>
- Bourrat, P. (2015). How to read ‘Heritability’ in the recipe approach to natural selection. *British Journal for the Philosophy of Science*, 66(4), 883–903.
- Bourrat, P. (2019). Heritability, causal influence and locality. *Synthese*, 198(7), 6689–6715.
- Bourrat, P. (2020). Causation and single nucleotide polymorphism heritability. *Philosophy of Science*, 87(5), 1073–1083.
- Chabris, C. F., Lee, J. J., Cesarini, D., Benjamin, D. J., & Laibson, D. I. (2015). The fourth law of behavior genetics. *Current Directions in Psychological Science*, 24(4), 304–312. <https://doi.org/10.1177/0963721415580430>
- Daniels, M., Devlin, B., & Roeder, K. (1997). Of genes and IQ. In B. Devlin, S. E. Fienberg, D. P. Resnick, & K. Roeder (Eds.), *Intelligence, genes and success: Scientists respond to the bell curve*. New York: Springer.
- Domingue, B. W., Belsky, D. W., Fletcher, J. M., Conley, D., Boardman, J. D., & Harris, K. M. (2018). The social genome of friends and schoolmates in the National Longitudinal Study of Adolescent to Adult Health. *Proceedings of the National Academy of Sciences*, 115(4), 702–707. <https://doi.org/10.1073/pnas.1711803115>
- Dowens, S. M., & Lucas, M. (2020). Heritability. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*.
- Duncan, L. E., & Keller, M. C. (2011). A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *The American Journal of Psychiatry*, 168(10), 1041–1049. <https://doi.org/10.1176/appi.ajp.2011.11020191>
- Falconer, D. S., & Mackey, T. F. (1996). *Introduction to quantitative genetics* (4th ed.). Essex: Longman.
- Feldman, M. W., & Lewontin, R. C. (1975). The heritability hang-up. *Science*, 190(4220), 1163–1168. <https://doi.org/10.1126/science.1198102>
- Galton, F. (1894). *Natural inheritance*. Macmillan and Company.
- Gauderman, W. J., Mukherjee, B., Aschard, H., Hsu, L., Lewinger, J. P., Patel, C. J., & Chatterjee, N. (2017). Update on the state of the science for analytical methods for gene-environment interactions. *American Journal of Epidemiology*, 186(7), 762–770. <https://doi.org/10.1093/aje/kwx228>
- Godfrey-Smith, P. (2009). *Darwinian populations and natural selection*. New York: Oxford University Press.
- Griffiths, A. J. F., Susan, R. W., Lewontin, R. C., Gelbart, W. M., Suzuki, D. T., & Miller, J. H. (2005). *Introduction to genetic analysis* (8th ed.). New York: W. H. Freeman.
- Haldane, J. B. S. (1938). *Heredity and politics*. London: George Allen & Unwin.
- Hartl, D. L., & Clark, A. G. (1997). *Principles of population genetics* (3rd ed.). Sunderland, MA: Sinauer.
- Hur, Y. M., & Bates, T. (2019). Genetic and environmental influences on cognitive abilities in extreme poverty. *Twin Research and Human Genetics*, 22(5), 297–301. <https://doi.org/10.1017/thg.2019.92>
- Jacquard, A. (1983). Heritability: One word, three concepts. *Biometrics*, 39(2), 465–477.
- Jencks, C. (1980). Heredity, environment, and public policy reconsidered. *American Sociological Review*, 45(5), 723–736. <https://doi.org/10.2307/2094892>
- Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., & Michelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic.
- Kaplan, J. M., & Turkheimer, E. (2021). Galton's quincunx: Probabilistic causation in developmental behavior genetics. *Studies in History and Philosophy of Science*, 88, 60–69. <https://doi.org/10.1016/j.shpsa.2021.04.001>
- Keller, E. F. (2013). Genes as difference makers. In K. Sheldon, & G. Jeremy (Eds.), *Genetic explanations: Sense and nonsense* (pp. 34–42). Harvard University Press.
- Kendler, K. S., & Baker, J. H. (2007). Genetic influences on measures of the environment: A systematic review. *Psychological Medicine*, 37(5), 615–626. <https://doi.org/10.1017/s0033291706009524>
- Knopik, V. S., Neiderhiser, J. M., DeFries, J. C., & Plomin, R. (2017). *Behavioral genetics* (7th ed.) New York.
- Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsdottir, B. J., Young, A. I., Thorgerisson, T. E., & Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science*, 359(6374), 424–428. <https://doi.org/10.1126/science.aan6877>
- Krapohl, E., & Plomin, R. (2016). Genetic link between family socioeconomic status and children's educational achievement estimated from genome-wide SNPs. *Molecular Psychiatry*, 21(3), 437–443. <https://doi.org/10.1038/mp.2015.2>
- Lee, P. N., Forey, B. A., & Coombs, K. J. (2012). Systematic review with meta-analysis of the epidemiological evidence in the 1900s relating smoking to lung cancer. *BMC Cancer*, 12, 385. <https://doi.org/10.1186/1471-2407-12-385>
- Lewontin, R. C. (1974). The analysis of variance and the analysis of causes. *American Journal of Human Genetics*, 26(3), 400–411.
- Lewontin, R. C. (2006). The analysis of variance and the analysis of causes. 1974. *Int. J. Epidemiol.*, 35(3), 520–525. <https://doi.org/10.1093/ije/dyl062>

- Lynch, K. E., & Bourrat, P. (2017). Interpreting heritability causally. *Philosophy of Science*, 84(1), 14–34.
- McGue, M., & Carey, B. E. (2017). Gene-environment interaction in the behavioral sciences: Findings, challenges, and prospects. In P. H. Tolan, & B. L. Leventhal (Eds.), *Gene-environment transactions in developmental psychopathology: The role in intervention research* (pp. 35–57). Cham, Switzerland: Springer International Publishing.
- Menzies, P., & Beebe, H. (2020). Counterfactual theories of causation. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <https://plato.stanford.edu/archives/win2020/entries/causation-counterfactual/>.
- Nelkin, D., & Andrews, L. (1996). The bell curve: A statement. *Science*, 271, 13–14.
- Nuffield. (2002). *Genetics and human behavior: The ethical context*. Retrieved from London.
- Oftedal, G. (2005). Heritability and genetic causation. *Philosophy of Science*, 72(5), 699–709. <https://doi.org/10.1086/508126>
- Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., & LifeLines Cohort, S. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics*. <https://doi.org/10.1038/s41588-022-01016-z>
- Pearson, C. H. (2007). Is heritability explanatorily useful? *Studies in History and Philosophy of Biological and Biomedical Sciences*, 38(1), 270–288. <https://doi.org/10.1016/j.shpsc.2006.12.012>
- Peñaherrera-Aguirre, M., Woodley, M. A., Sarraf, M. A., & Beaver, K. M. (2022). Social adversity reduces polygenic score expressivity for general cognitive ability, but not height. *Twin Research and Human Genetics*, 25(1), 10–23. <https://doi.org/10.1017/thg.2022.3>
- Pesch, B., Kendzia, B., Gustavsson, P., Jöckel, K. H., Johnen, G., Pohlbeln, H., & Brüning, T. (2012). Cigarette smoking and lung cancer—relative risk estimates for the major histological types from a pooled analysis of case-control studies. *International Journal of Cancer*, 131(5), 1210–1219. <https://doi.org/10.1002/ijc.27339>
- Plomin, R. (2014). Genotype-environment correlation in the era of DNA. *Behavior Genetics*, 44(6), 629–638. <https://doi.org/10.1007/s10519-014-9673-7>
- Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 11(1), 3–23. <https://doi.org/10.1177/1745691615617439>
- Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2008). *Behavioural genetics* (5th ed.). New York: Worth.
- Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, 47(7), 702–709. <https://doi.org/10.1038/ng.3285>
- Rutter, M. (1997). Nature-nurture integration: The example of antisocial behavior. *American Psychologist*, 52, 390–398.
- Rutter, M. (2002). Nature, nurture, and development: From evangelism through science toward policy and practice. *Child Development*, 73, 1–21.
- Sasco, A. J., Secretan, M. B., & Straif, K. (2004). Tobacco smoking and cancer: A brief review of recent epidemiological evidence. *Lung Cancer*, 45(Suppl. 2), S3–S9. <https://doi.org/10.1016/j.lungcan.2004.07.998>
- Schaffer, J. (2016). The metaphysics of causation. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2016 Edition). <https://plato.stanford.edu/archives/fall2016/entries/causation-metaphysics/>.
- Sesardic, N. (2005). *Making sense of heritability*. New York: Cambridge University Press.
- Sober, E. (2001). Separating nature and nurture. In R. W. D. Wasserman (Ed.), *Genetics and criminal behaviour* (pp. 47–78). Cambridge: Cambridge University Press.
- Tal, O. (2009). From heritability to probability. *Biology and Philosophy*, 24(1), 81–105.
- Tucker-Drob, E. M., & Bates, T. C. (2016). Large cross-national differences in gene × socioeconomic status interaction on intelligence. *Psychological Science*, 27(2), 138–149. <https://doi.org/10.1177/0956797615612727>
- Turkheimer, E. (2016). Weak genetic explanation 20 years later: Reply to Plomin et al. (2016). *Perspectives on Psychological Science*, 11(1), 24–28. <https://doi.org/10.1177/1745691615617442>. <https://plato.stanford.edu/archives/fall2016/entries/causation-metaphysics/>.
- Turkheimer, E., Goldsmith, H. H., & Gottesman, I. I. (1995). Commentary. *Human Development*, 38(3), 142–153. <https://doi.org/10.1159/000278307>
- West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford: Oxford University Press.
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.
- Woodward, J. (2016). Causation and manipulability. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from London.
- Yang, J., Zeng, J., Goddard, M. E., Wray, N. R., & Visscher, P. M. (2017). Concepts, estimation and interpretation of SNP-based heritability. *Nature Genetics*, 49(9), 1304–1310. <https://doi.org/10.1038/ng.3941>
- Young, A. L., Benonisdotir, S., Przeworski, M., & Kong, A. (2019). Deconstructing the sources of genotype-phenotype associations in humans. *Science* (New York, N.Y.), 365(6460), 1396–1400. <https://doi.org/10.1126/science.aax3710>