Philosophy 3334: Philosophy of Biology
Fall 2023 - Homework 5 - answers

Answers should be uploaded into Blackboard before 11:59pm on Tuesday, Nov 21.

1) Let's assume a farmer has two genetic varieties (G_1 and G_2) of the same type of pumpkin. They plant them in four different fields (E_1 – E_4) where each field has one of two different types of soil (S_1 or S_2). Then one field of each type of soil is fertilized and the other is not. One hundred of each type of pumpkin are planted in each field. The average mass in kilograms of each type of plant in each field is given in the chart below:

	G_1	G_2	average
E_1 (S1 + F)	4	12	8
E_2 (S1 without F)	2	8	5
E_3 (S2 + F)	8	8	8
E_4 (S2 without F)	2	4	3
average	4	8	6

Overall mean = 6

From this information, calculate each of the following quantities (and show your work):

For these problems we will need averages (means). The overall mean is M = 48/8=6. M(g1) = 4, M(g2) = 8, M(e1) = 8, M(e2) = 5, M(e3) = 8, M(e4) = 3. I have added those numbers in the chart above.

Vp – overall phenotypic variance
This is the overall 'spread' of the data. What is the average (squared) distance from the overall mean. Formally, $Vp=1/n \Sigma (x-m)^2$. In this case that is:

$Vp = 1/8\ [(4-6)^2 + (12-6)^2 + (2-6)^2 + (8-6)^2 + (8-6)^2 + (8-6)^2 + (2-6)^2 + (4-6)^2] = 11.$

Vg – overall genotypic variance
Vg = treats each of the genotypes as a single thing with its mean for its phenotype. Then it looks at the spread around the overall mean. So for example, if the genotypes are the same *on average* then Vg overall is 0. In this case,

$Vg = ½\ [(M(g1) - M)^2 + (M(g2)-M)^2] = ½\ [(4-6)^2 + (8-6)^2] = 4$

Ve – overall environmental variance
$Ve = 1/4\ [(M(e1) - M)^2 + (M(e2)-M)^2 + (M(e3) -M)^2 + (M(e4)-M)^2] =$
$= 1/4\ [(8-6)^2 + (5-6)^2 + (8-6)^2 + (3-6)^2] = 18/4 = 4.5$

Vgxe – overall variance due to the interaction of genetic and environmental factors (called I in Sober). NOTE: In this case, because there are the same number of each type of plant in each type of environment the covariance between genotype and environment Cov(g,e) = 0 so you do not need to worry about this term.

In full generality, Vp = Vg + Ve + Vgxe + 2Cov(g,e). Covariance of g and e refers to how much more likely some genes are to be found in some environments rather than others. In this case, there are 100 plants in each type of environment so the covariance is 0. In these kinds of cases, Vp = Vg + Ve + Vgxe. Therefore in this case, Vgxe = Vp-Vg-Ve = 2.5

H^2 - broad score heritability (h^2 in Sober)

H^2 = Vg/Vp. This is the 'percent of the variation explained by genetic variation'. In this case, H^2 = 4/11.

2) Now let's use the data from problem 1 to ask about causation. Calculate the following:

2a) The average mass of a pumpkin in fertilized soil.
This is the average of pumpkins in E1 and E3 = 8
2b) The average mass of a pumpkin in non-fertilized soil.
This is the average of pumpkins in E2 and E4 = 4
2c) The average mass of a pumpkin in soil type 1.
This is the average of pumpkins in E1 and E2 = 6.5
2d) The average mass of a pumpkin in soil type 2.
This is the average of pumpkins in E3 and E4 = 5.5

2e) Harden doesn't actually give a definition of causation in her book, but she does say on page 108, "All" that is required to assert that you have identified a cause is to demonstrate evidence that the average outcome for a group of people *would have been different* if they had experienced **X** instead of **Not-X**. – According to this, do you think Harden would say that the fertilizer causes the pumpkins to be larger? What about being in soil type 1? Carefully explain your answers. Do you think this is correctly capturing causation here? What would happen if almost all of the pumpkins of this type in the wild were G_1 and G_2 was a rare mutant type? Would this change the answer at all? Should it?

In asking whether the fertilizer is a cause, we want to know of some particular group of plants, whether they *would have been different* if they had experienced the fertilizer instead of not receiving the fertilizer. In this case, it is natural to take the group of plants that were not fertilized (average height of 4) and ask what would have happened if they had been fertilized. Some have genotype 1, some g2, some are in E1, some in E2, etc. But each group has a natural group of plants to compare to that they are like in all respects other than being fertilized so it is natural to think that the newly fertilized plants would be just like the plants that were actually fertilized. So in that case, their average mass would have been 8 which is larger than 4. So yes, there is evidence that fertilizer is a cause of having larger mass. The same reasoning applies to soil type 1 being a cause relative to soil type 2 (where soil type 2 is 'not X'). So Harden should say that soil type 1 is a cause of larger mass as well.

However, it is not clear that this is correct. What is actually happening is that soil type 1 is good for plants of type G2 whereas soil type 2 is better for plants of type G1. However,

the gain for G2 in S1 is larger than the gain for G1 in S2 so on average, it looks like S1 is better. But if we ask whether some group of plants would have been larger if they experienced S1, it makes a difference which kind of plants we started with. So if we were looking at all the pumpkins in the wild and almost all of them were of type G1, then actually, S2 would be better on average.

To get full credit for this answer you need to show some recognition that S1 is better for one subgroup and worse for the other. But you could say all that matters is the average (and then causation is population relative) or you could say that it is inappropriate to say S1 causes larger mass *simpliciter* but that it should only be a cause if it is a cause for all the plants in the group.

3) On page 121 Harden says, "Even as I write this, I hear a chorus singing out a familiar objection: 'Heritability estimates are specific to a population.' "

Imagine that the farmer does the exact calculation you did in problem #1 and says, "Now I have learned something interesting about pumpkins. I have learned that the heritability of size in these pumpkins is xxx [your answer]. That is really interesting and valuable information." Now a critic comes along and says, "Who cares? This is worthless information. Don't you know that heritability estimates are specific to a population?"

Explain this objection. What is the critic trying to tell the farmer here?

When we calculated that H^2 was 4/11 in this population of pumpkins, that particular answer is relative to these pumpkins in these exact environmental conditions. For example, if next year the farmer fertilized all of their plants or put them all in soil type 1, the heritability value would be different. So knowing that it is 4/11 for this particular population is not 'portable' (Harden's term for applicable in other situations) and so not valuable.

NOTE: Harden acknowledges that heritability is population relative, but believes that it is still very valuable information nonetheless because she thinks it is telling us what explains the phenotypic differences in particular populations. So if the farmer wanted to know why these particular pumpkins were much larger than these other pumpkins in what be appropriate to say that about 36% of the variation is explained by genetic differences while about 41% is explained by environmental differences.

4) Some traits in humans seem to be highly heritable – for example, height is around $H^2 = .8$ as is schizophrenia. To many people this suggests that your genes almost completely determine your height and similarly, nothing we do will affect whether or not a child will develop schizophrenia later in life. This would seem even more powerful if the trait had a heritability value of .95 or even 1. But actually, this is incorrect.

4a) Explain how this high value for heritability is consistent with the claim that a new, different environment might affect phenotypes.

High heritability of a trait in a population simply means that there are genetic differences between individuals that are correlated with phenotypic differences in that group and actual environmental differences are not correlated with phenotypes. If we introduced a new environment, that new environment (unlike the old environments) may affect some individuals much more than others.

4b) Give an example where this has happened in the past (or for fun, describe an example (realistic sounding or not) of an alteration like this that could happen in the future).

Sober and Harden both mention eyeglasses and phenylketonuria (PKU) as cases where a trait (vision or PKU disease) is highly heritable because in the past, these phenotypes were highly correlated with particular genetic bases. But of course eyeglasses affect vision and new diets low in phenylalanine can prevent the disease.

Taking an example like schizophrenia to be highly heritable it is easy to imagine a pill or brain surgery that prevents the relevant phenotype. For something like eye color or blood type it is harder to imagine, but again, some kind of medical intervention could potentially break the connection between genotype and phenotype.

4c) Explain how things actually happening in the environment right now could still be causally influencing these traits.

Heritability is capturing whether differences in actual environments have differential effects on phenotypes. So as long as all the different actual environments are affecting the phenotypes the same (on average) it will not show up as part of Ve. So for example, if everyone is eating a diet rich in phenylalanine, this is causally leading to PKU disease but doesn't show up because everyone has the same diet. Maybe minerals in the soil are affecting plant growth or pollution in the air is giving us cancer but removing it would help. Similarly for mundane things like oxygen in the air or drinking water whether or not we could actually alter this without dying.

5) One common kind of twin study involves comparing how similar the phenotypes of identical twins are to how similar the phenotypes of fraternal twins are. The standard methodology assumes what is often called the "equal environments assumption."

5a) Explain what this assumption means.
5b) Explain why it is almost certainly false.
5c) Explain how this assumption matters to the calculation of heritability (either formally or informally) and what the effects are if it is false.

The “equal environments assumption” says that however similar the common environment of identical twins raised together is, the environments are just as similar (or dissimilar) as the common environments of fraternal twins raised together. Formally, we assume Ve(mono) = Ve(diz). This is probably false. Parents, friends, etc. shape the environments around their children partly in virtue of traits the children have. So if they are athletic and strong and show an interest in sports, they might join a sports team or get lots of coaching and encouragement. Similarly for musical talent or an interest in art. If they are sick a lot they might stay home from school a lot or receive different kinds of parenting. But identical twins are more similar in these ways than fraternal twins are and so probably their environments are more similar as well.

If Ve(mono) < Ve(diz) then the differences between the two types of twins are not all genetically caused and so the genes are really causing/explaining less of the differences and so h^2 is being overestimated. It should be lower. Mathematically, looking at Sober’s paper, if Ve(mono) < Ve(diz) then (17) is wrong then (18) is wrong because Vp(diz) is actually higher than the right-hand-side so in (19) the left hand side is greater so in (20) the left hand side is less. So Vg(diz) is actually smaller than we previously calculated. So by 21, Vg(everyone) is actually smaller than we calculated. So $Vg/Vp = h^2$ is smaller than we calculated.