



Considering Instructional Approach & Question Design with the Hardy-Weinberg Principle

Authors: Driver, Robert J., and McRae, Susan B.

Source: The American Biology Teacher, 83(3) : 191-194

Published By: National Association of Biology Teachers

URL: <https://doi.org/10.1525/abt.2021.83.3.191>

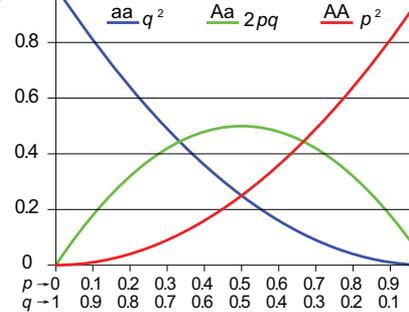
BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

ROBERT J. DRIVER, SUSAN B. McRAE

**ABSTRACT**

The Hardy-Weinberg principle (HWP) is an application of the binomial expansion theorem that is foundational to the field of population genetics. Because of the important history of the HWP in answering how variation is preserved during evolution, and the ability of Hardy-Weinberg equilibrium (HWE) to detect natural and sexual selection acting on a trait, the HWP is a staple of the introductory biology undergraduate curriculum in the United States. Introductory courses often cover a wide range of topics in ecology and evolution, and it is important that students have enough time during the semester to grasp the foundations of population genetics. At the same time, information needs to be presented clearly to ensure that the student gains a correct understanding of the HWP. This article discusses the importance of the HWP to undergraduate education in biology and describes misconceptions from the instructor's perspective. These misconceptions are pervasive and risk undermining a proper understanding of the HWP. We provide examples adapted from university- and AP-level standardized tests.

Key Words: Hardy-Weinberg; population genetics; introductory biology; mathematics; evolution.

○ The HWP: Historical Perspective

A critical link between Darwin's theory of evolution and Mendelian genetics was provided by G. H. Hardy and Wilhelm Weinberg's analyses of the segregation of trait variants following Mendelian inheritance rules (Hardy, 1908; Weinberg, 1908), contributing what is known as the Hardy-Weinberg principle (HWP). Darwin's theory had not fully accounted for the preservation of diversity through direct inheritance (Vorzimmer, 1968). "Blending inheritance," a leading theory in the late 19th century, attacked Darwinian evolution by pointing out the impossibility of preserving diversity if traits inherited from each parent led to intermediate offspring (Jenkin, 1867). The rediscovery of Mendel's findings illustrated an inheritance pattern where parental phenotypes were preserved through generations without necessarily being expressed (Mendel, 1866; Druery & Bateson, 1901). The HWP, a form of the binomial

expansion theorem, provided the mathematical rationale for Mendelian inheritance, rendering blending inheritance obsolete.

In its simplest form, Hardy-Weinberg equilibrium (HWE) assumes a Mendelian inherited gene with only two alleles *A* and *a*, with respective frequencies *p* and *q*. Given only two variants in the population, these proportions must add to one ($p + q = 1$). Two binomials are generated, representing the probability of inheriting either allele from the father, $(p_m + q_m)$, and from the mother, $(p_f + q_f)$. The product of the probabilities, $(p_m + q_m)(p_f + q_f)$, analogous to the binomial probability of flipping a coin twice in succession, equates to one representing the whole population. Multiplying the terms ("FOIL" mnemonic) then gives the Hardy-Weinberg equation, $p^2 + 2pq + q^2 = 1$, where the frequency of receiving two *A* alleles is p^2 , two *a* alleles is q^2 , and one of each is $2pq$ (the heterozygote genotype occurs twice, representing alternative inheritance of each allele from each parent, $p_m q_f$ and $p_f q_m$). Provided that allele frequencies do not change between generations (HWE conditions), the genotypic and phenotypic frequencies remain the same. This resolves Darwin's predicament and explains why diversity is maintained across generations.

○ Present-Day Context of the HWP in the Undergraduate Classroom

Appropriately, the HWP is a gateway into the application of evolutionary theory to populations in biology undergraduate curricula (Journet, 1986). The HWP helps students understand the connection between evolution in a single pedigree and evolution in a population, highlighting changes in allele frequencies over time (Mertens, 1992). Once the background and significance of the HWP is presented, the math behind the HWP is typically introduced. In HWP math problems, students are often given the number of homozygous and heterozygous individuals in a population and asked to calculate allele frequencies to determine whether the original genotypic frequencies provided match genotypic frequencies generated from the Hardy-Weinberg equation. To determine whether the genotypic frequencies from the given population are statistically significant from the genotypic frequencies at HWE, students may

employ a chi-square test and compare values. This also serves as an introduction to biological statistics for undergraduates.

Upon calculating a p -value from the chi-square test, students will find that the population either matches the values from the Hardy-Weinberg equation (is in HWE) or does not match the genotypic frequencies from the equation (is not in HWE). As a final step, students then relate their findings back to the biology of the population. If the genotypic frequencies in question do not match HWE, what does that tell us about the population with respect to that locus? The HWP describes how variation persists when allele frequencies remain constant (population is not evolving with respect to the trait). In natural populations evolving with respect to a trait, allele frequencies change over time. To understand why a locus may not be in HWE, students learn about the assumptions of the HWP. HWE requires populations to have a very large size, so as to negate the effects of genetic drift. At HWE, neither selection nor nonrandom mating can give one genotype a fitness advantage over the other. Finally, no migration into the population and no mutation can impact the neutral attributes of either allele. Students can hypothesize about which of these forces may be in play in a given example.

Despite the importance of the HWP, students may have difficulty solving problems and grasping the connection between the math and the biology. The math itself can be intimidating to students, and common pitfalls can be avoided with innovative teaching techniques (Ortiz et al., 2000; Masel, 2012; Brewer & Gardner, 2013). However, to make sure that students gain a thorough understanding of the HWP, biology instructors must also be adept at working through problems and addressing student concerns. Unfortunately, common misconceptions are propagated through instructor error (even at the university level) as well as on standardized tests. Given the fundamental nature of the HWP in population genetics and evolution generally, failing to teach the HWP correctly can lead to long-lasting confusion for biology majors, particularly with concepts that build upon the implications of the HWP. Below, we address misconceptions derived from teaching materials and questions from standardized tests, and suggest how to present information to students so as to build an understanding of evolution in populations.

○ Issues & Solutions

Solution 1: Make Sure Equilibrium Is Not Stated

One important goal for students learning the HWP is to understand how an evolving population differs from one under HWE. This concept is conveyed to students in part by determining whether or not a given population is in HWE. Therefore, it is important that sample problems do not specify whether a population is in HWE, but to allow the student to come to this conclusion themselves (Figure 1). The decision of whether the population mathematically meets HWE conditions allows the student to connect a biological interpretation of evolution with statistical methods. By contrast, in problems where HWE is assumed, the count of each genotype and genotypic frequencies of the population are often not given, thereby preventing the possibility of working through the problem with a biological context. Particularly at issue is that questions specifying populations at HWE enable students to find allele frequencies by taking square roots of homozygote genotype frequencies. This shortcut backfires on problems where given allelic frequencies are not in HWE, thereby confusing the student by introducing an

Researchers discover that a recessive allele causes heightened hearing and the ability to detect low-frequency sounds (*hear* allele). In the general population, the phenotype occurs in 1 in 100,000 live births. In an isolated population, however, the phenotype occurs in 1 in 300 live births.

Assume that both the isolated and general populations are in Hardy-Weinberg equilibrium (HWE) for this locus. What is the allele frequency of the *hear* allele in both populations?

Answer: The frequency of the *hear* allele is the square root of $1/300$ or 0.06 in the isolated population and the square root of $1/100,000$ or 0.003 in the general population.

Alternative problem:

A recessive allele (*hear*) causes heightened hearing and the ability to detect low-frequency sounds. In the general population, based on a very large sample, the phenotype occurs in 1 in 100,000 live births. A PCR diagnostic test is developed to detect the allele. A screened sample of 1000 unrelated individuals from a remote island are given the test: 995 are homozygous dominant, 1 is homozygous recessive, and 4 are heterozygote. Is this trait under selection?

Answer:

Under HWE, the allele frequencies of the island population can be computed using $p = 0.995 + \frac{1}{2}(0.004) = 0.997$, and $q = 0.001 + \frac{1}{2}(0.004) = 0.003$. Then, the predicted population frequency of the recessive allele under HWE would be $q^2 = 0.000009$ or 0.9 in 100,000. To the nearest individual, based on a χ^2 test, this ratio is equivalent to 1 in 100,000, so the locus does not appear to be under selection.

Figure 1. Example of an HWP problem where HWE is specified a priori. The question does not allow students to test evolution at this locus. The question does not present the number of individuals in the two populations for each genotype, so the only genotypic frequencies capable of being calculated are through the Hardy-Weinberg equation. The student can manipulate variables in the equation, and convert genotypic frequencies to allelic frequencies, but by assuming that both populations are in HWE, the problem is not representative of a challenge a researcher would naturally encounter. In our alternative example, we provide a sample population for students to test, and we do not specify HWE. The locus here is not under selection, but we have allowed the student to work through the problem to come to this conclusion. (Example adapted from a question in the 2013 AP Biology Practice Exam from College Board.)

entirely different set of methods than what are necessary for solving a biologically meaningful HWP problem (Figure 2).

Solution 2: Genotype Counts (or Frequencies) Must Be Provided if HWE Is Not Specified

We do not recommend that an instructor use a problem where HWE is assumed, but if this does occur, then genotype counts must be provided (Figure 3). This is because a problem that does not specify HWE and also does not specify genotype counts or frequencies cannot be solved. Students need to compare the genotype frequencies in the sample population to a population in HWE. If a problem specifies HWE, but does not give genotype counts, then students can still conclude that the population is in HWE, because this information is given. However, if students are not given genotype counts or information about HWE, then the problem cannot be solved. Students can try to solve the problem using incorrect means by either assuming that the population is in HWE or inventing their own genotype counts for the population, in which case answers for each student would be different. In these problems that do not specify HWE or genotype counts, the allelic frequencies p and q are provided, allowing the student to figure out the genotypic frequencies if the population were in HWE, through the Hardy-Weinberg equation.

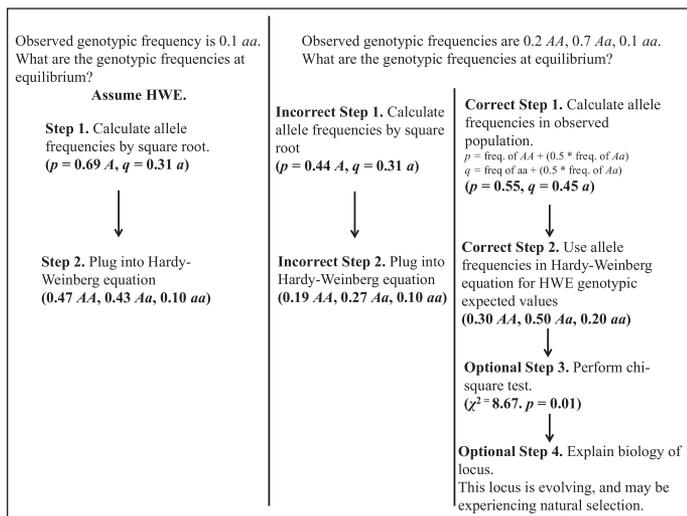


Figure 2. Schematic for solving HWP questions. The different steps required to solve a problem that assumes HWE and a problem that does not are shown. Problems that assume HWE are often solved by taking the square root of a given homozygotic genotypic frequency to get the allele frequency (left column). However, taking the square root of a given homozygotic genotypic frequency when the problem does not specify HWE leads to an incorrect answer (center column). Instead, if all problems do not specify HWE, students can learn a consistent procedure for testing evolutionary theory across all problems (right column).

Asking the student to *assume* that HWE is met encourages the student to always make this assumption, even when it is incorrect to do so in biologically meaningful HWP problems. These problems may also ask the question of how the HWP can be used to determine whether the example population is evolving (Figure 3). It is only possible to obtain expected genotypic frequencies assuming HWE, so the student will have nothing to compare these frequencies to. If taught to use this incorrect strategy to solve this type of problem, the student is also likely to use the same strategy in HWP problems with a biological context and conclude erroneously that the genotypic frequencies generated by the Hardy-Weinberg equation are the only values relevant to this concept. This undermines efforts to correctly work through an HWP problem.

Solution 3: Be Sure That Wording Accurately Reflects the HWP

To gain the interest of students pursuing medicine, HWP examples often use genetic conditions with deleterious alleles (Figure 4). Although biological information about the relevant locus in an HWP problem is superfluous to solving the problem, it is beneficial to include when students consider HWE assumptions that may be violated. Therefore, we recommend adding information about the locus that an HWP problem is focused on, as it helps relate the question to the real world. Language can describe the phenotypic effects or the cellular pathways that the protein functions in. However, it is important to consider that alleles specifically described to have major fitness costs should not, in turn, be defined as experiencing HWE. If the student is thinking about the biology of the problem, an allele causing cancer (Figure 4) should present a selective disadvantage and not be in HWE. Additionally, it is important

Evolution involves changes in allele frequencies in a population over generations. For a particular genetic locus in a population, the frequency of the recessive allele (*b*) is 0.3.

(a) What is the frequency of each of the genotypes (*BB*, *Bb*, *bb*) in this population? What is the frequency of the dominant phenotype?

(b) How can the Hardy-Weinberg principle be used to determine whether the population is evolving?

(c) Identify a violation of Hardy-Weinberg equilibrium (HWE) and how it might alter allele frequencies in the population.

Corrected problem:

Evolution involves changes in allele frequencies in a population over generations. **Whereas, when neither drift nor selection is acting upon a specified allele, the allele and genotype frequencies remain in equilibrium across generations.** For a particular genetic locus with **only two alleles**, the frequency of the recessive allele (*b*) is 0.3.

(a) **Assuming that the locus is under HWE, what is the predicted** frequency of each of the genotypes (*BB*, *Bb*, *bb*) in this population? What is the **predicted** frequency of the dominant phenotype?

(b) How can the Hardy-Weinberg principle be used to determine whether the population is evolving?

(c) **Could deviations from the Hardy-Weinberg equilibrium reveal whether an allele may be advantageous or disadvantageous? Explain.**

Figure 3. Example of a problem where genotype counts are not given and HWE is not specified. Here, it is likely that instructors intend for students to assume HWE, but the assumption of HWE is not specified. Therefore, students can plug the given allele frequencies into the Hardy-Weinberg equation. The question fails to reinforce concepts of evolution. If students assume HWE, they practice an incorrect approach to future problems where HWE is not specified and the population is not in equilibrium. In our corrected problem, we provide definitions for loci subjected to evolution and for loci under HWE conditions. We then specify HWE in the problem, and finally connect the locus in question to advantages or disadvantages in evolution. By asking students about how HWE relates to evolution, even though HWE is specified here, we still challenge the student to discuss the topic. (Example adapted from the 2008 AP Biology Exam Free Response Questions, from College Board.)

to remember that the HWP works independently on each locus across the genome, and so different loci will vary in the degree to which they violate the HWP. Problems should ask whether a single locus is in HWE, as opposed to populations (Figure 4). Describing populations as under HWE takes away from the student's understanding that the HWP can be applied independently to any locus and gives the impression that all loci behave in the same way.

○ Conclusions

The HWP enables us to understand how diversity is preserved and how evolution operates on allelic frequencies over the course of generations. The concept allows students to think about how loci violate HWE assumptions and enables them to think creatively when reconciling the impact of genotype on phenotype. It also introduces the application of statistics in a biological framework. We describe three solutions to misconceptions of the HWP and provide examples of each adapted from standardized tests. We encourage faculty and graduate students teaching introductory biology courses to seek out population geneticists or evolutionary biologists in their departments to help ensure that all instructors leading students through the HWP understand the concept properly and

In a population, 1 out of every 500 people has a cancer caused by a completely recessive allele, *c*. Assuming that the population is in Hardy-Weinberg equilibrium (HWE), what is the expected proportion of individuals in the population heterozygous for the cancerous *c* allele?

Answer: About 43 individuals in the population of 500 individuals

Alternative problem:

In a population, 1 out of every 500 people **test positive as homozygotes for a completely recessive allele (*c*)**. Assuming that the population is in HWE for this locus, what is the expected proportion of individuals in the population heterozygous for **the *c* allele?**

Answer: About 43 individuals in the population of 500 individuals

Better problem showing that particular genotypes tend to be risk factors for disease:

About 25% of patients diagnosed with a particular form of cancer are found to be homozygous at a locus with only two alleles for the recessive allele, *c*. The rate of this cancer in the general population is 1 out of every 500 people. To determine whether there is selection on this allele, genotypes of 10,000 people are screened at this locus based on voluntarily submitted genome data. Only 99 are found to have the homozygous recessive genotype, but 4152 are carriers of one *c* allele. Has there been selection against this allele? Interpret the observed pattern of genotypes.

Answer:

Use the observed genotypes to obtain allele frequencies: $q = 0.0099 + \frac{1}{2}(0.4152) = 0.2175$; $p = 0.5749 + \frac{1}{2}(0.4152) = 0.7825$. Then, the predicted population frequency of the recessive allele under HWE would be $q^2 = 0.0473$ or 473 out of 10,000 individuals (or 4.73%), and the predicted proportion of heterozygotes $2pq = 0.3404$ or 3404 out of 10,000. A χ^2 test would determine that these predicted genotype frequencies are significantly different from the observed. The homozygous recessive condition is less common than expected according to the HWP, suggesting that, yes, there may be selection against this allele. Also, the observed and expected frequencies of homozygous recessives exceed the observed cancer frequency of 1 in 500 (0.2%). Thus, homozygous recessives appear to be at a higher risk for the cancer (25% of cancer patients compared to 0.99% of the population), but they are not guaranteed to be afflicted.

Figure 4. Example of a problem featuring a recessive deleterious allele, but where HWE is specified. The population is defined as in HWE, when the question is only describing a single locus. This example prevents the student from understanding what HWE means in a population genetics context, forcing the student to associate HWE with one of its violations. By applying HWE to the population, and not a locus, students are not presented with a realistic description of how genomes evolve. In our improved problem, we provide a detailed explanation of the scenario, do not specify HWE due to the deleterious nature of the allele, and refer to testing the locus for HWE, not the population. (Example adapted from the 2017 GRE Biology Test Practice Book from ETS.)

apply best practices in teaching it. We hope that misleading HWP problems will be removed from standardized tests in the future and be replaced by ones that enhance student understanding.

References

- Brewer, M.S. & Gardner, G.E. (2013). Teaching evolution through the Hardy-Weinberg principle: a real-time, active-learning exercise using classroom response devices. *American Biology Teacher*, 75, 476–479.
- Druery, C.T. & Bateson, W. (1901). Experiments in plant hybridization. *Journal of the Royal Horticultural Society*, 26, 1–32.
- Hardy, G.H. (1908). Mendelian proportions in a mixed population. *Science*, 28, 49–50.
- Jenkin, F. (1867). Review of 'The Origin of Species.' *North British Review*, 46, 277–318.
- Journet, A.R.P. (1986). Population genetics: a fishy process. *American Biology Teacher*, 48, 478–482.
- Masel, J. (2012). Rethinking Hardy-Weinberg and genetic drift in undergraduate biology. *Bioessays*, 34, 701–710.
- Mendel, J.G. (1866). Versuche über Pflanzenhybriden. Verhandlungen des naturforschenden Vereines in Brünn, Bd. IV für das Jahr. *Abhandlungen*, 3–47.
- Mertens, T.R. (1992). Introducing students to population genetics & the Hardy-Weinberg principle. *American Biology Teacher*, 54, 103–107.
- Ortiz, M.T., Taras, L. & Stacoulakis, A.M. (2000). The Hardy-Weinberg equilibrium: some helpful suggestions. *American Biology Teacher*, 62, 20–22.
- Vorzimmer, P.J. (1968). Darwin and Mendel: the historical connection. *Isis*, 59, 77–82.
- Weinberg, W. (1908). Über den Nachweis der Vererbung beim Menschen. *Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg*, 64, 368–382.

ROBERT DRIVER (driver16@students.ecu.edu) is a graduate student and SUSAN B. McRAE is a Teaching Professor in the Department of Biology, East Carolina University, Greenville, NC 27858. Funding provided by National Science Foundation awards RCN-1457541 and IOS-1456612.