Journal of Biological Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rjbe20

The Conceptual Analysis of Disciplinary Evidence (CADE) framework as a guide for evidentiary reasoning: A practical implementation in a Hardy-Weinberg Equilibrium (HWE) laboratory investigation

Chaonan Liu, Dayna L. Dreger, Shiyao Liu, Ala Samarapungavan, Stephanie M. Gardner, Kari L. Clase & Nancy Pelaez

To cite this article: Chaonan Liu, Dayna L. Dreger, Shiyao Liu, Ala Samarapungavan, Stephanie M. Gardner, Kari L. Clase & Nancy Pelaez (2022): The Conceptual Analysis of Disciplinary Evidence (CADE) framework as a guide for evidentiary reasoning: A practical implementation in a Hardy-Weinberg Equilibrium (HWE) laboratory investigation, Journal of Biological Education, DOI: 10.1080/00219266.2022.2100450

To link to this article: https://doi.org/10.1080/00219266.2022.2100450

JOURNAL OF BIOLOGICAL EDUCATION
https://doi.org/10.1080/00219266.2022.2100450

The Conceptual Analysis of Disciplinary Evidence (CADE) framework as a guide for evidentiary reasoning: A practical implementation in a Hardy-Weinberg Equilibrium (HWE) laboratory investigation

Chaonan Liu [a], Dayna L. Dreger [b], Shiyao Liu [a], Ala Samarapungavan [c], Stephanie M. Gardner [a], Kari L. Clase [d] and Nancy Pelaez [a]

[a]Department of Biological Sciences, Purdue University, West Lafayette, IN, USA; [b]Department of Basic Medical Sciences, College of Veterinary Medicine, Purdue University, West Lafayette, IN, USA; [c]Department of Educational Studies, Purdue University, West Lafayette, IN, USA; [d]Department of Agricultural and Biological Engineering; Biotechnology Innovation Regulatory Science (BIRS) Center, Purdue University, West Lafayette, IN, USA

ABSTRACT

Recent emphasis on learning biology through scientific investigations has focused instruction on understanding and using scientific evidence. To unpack the complexities of evidentiary reasoning, here we present a novel laboratory investigation for teaching the Hardy Weinberg Equilibrium (HWE) in an introductory biology laboratory course that was informed by the Conceptual Analysis of Disciplinary Evidence (CADE) framework. This HWE laboratory investigation highlights evidentiary reasoning with scaffolding questions to target reasoning with and about scientific evidence and it provides students with detailed disciplinary knowledge underpinning their investigation and evidentiary reasoning. To indicate how CADE influenced instruction on evidentiary reasoning during the investigation, changes in one instructor's laboratory discussion questions before and after implementing CADE are provided. Also, our CADE-informed HWE laboratory investigation is compared to other published activities for engaging students with HWE. Findings show that with CADE, the instructor implemented more scaffolds for directing students to consider multiple aspects of evidentiary reasoning and encouraged students' epistemic considerations about the nature, scope and quality of scientific evidence. These changes suggest that CADE can be a practical pedagogical tool to inform improvements in HWE laboratory investigations and to help instructors develop and implement scaffolds to guide students' evidentiary reasoning during a scientific investigation.

KEYWORDS

Evidence; reasoning; scaffolding; Hardy-Weinberg equilibrium; undergraduate; laboratory course

Scientific evidence and evidentiary reasoning

As we enter the era of 'big data' and the rapid advancement of the digital world, people are constantly challenged to evaluate data as evidence and make informed decisions that impact both daily and professional life (Sarit and Chinn 2020; Labrinidis and Jagadish 2012; Marx 2013). For example, during the COVID-19 pandemic, information about the current situation was posted on TV shows, newspapers, social media, and reports from government officials. However, even information from the most widely trusted resource is prone to error and can be interpreted

CONTACT Chaonan Liu liucn89@gmail.com
Supplemental data for this article can be accessed online at https://doi.org/10.1080/00219266.2022.2100450.

differently by different people (Zarocostas 2020). Thus, educating the public and the future workforce how to properly evaluate evidence poses great challenge to tertiary level education (Sarit and Chinn 2020; Chinn, Barzilai, and Golan Duncan 2020). There is a need for studies on how to equip students at tertiary level, especially in science disciplines, with the fundamental competency of understanding and using scientific evidence, which will benefit their professional development as well as daily life.

From the perspective of science education, scientific evidence is based on data that are used to address a question or used in the process of supporting a claim (Next Generation Science Standards Lead States 2013; Sandoval and Reiser 2004). Scientific evidence in authentic research is usually generated with various technologies and methods that rely on multiple types of disciplinary knowledge and practices. In real investigations, the evidence for claims varies in amount, scope, and comprehensiveness, whereas scientific evidence used in the science classroom is often more simplified in forms and usages (Chinn and Malhotra 2002; Duncan, Chinn, and Barzilai 2018). During science learning, students rarely encounter the kinds of complex and more contentious evidence that scientists often encounter in their research. The gap that exists between the simplified evidence used in science education and the actual complexity of scientific evidence that exists in the real world may contribute to the difficulty for students to understand scientific evidence as well as to develop evidentiary reasoning ability, which refers to their competence for reasoning with and about scientific evidence in their science learning (Duncan, Chinn, and Barzilai 2018; Samarapungavan 2018). Helping students properly understand, use and reason with scientific evidence in the process of an investigation also presents a challenge for the instructor who guides the investigations.

Despite numerous reports on attempts to improve students' evidentiary reasoning ability, studies have revealed that students still have problems with understanding and using evidence. Sandoval and Millwood (2005) showed that middle school students often failed to cite sufficient evidence when writing explanations for natural selection problems, and the students did not articulate the connection between the specific evidence and the claims. In the context of atmospheric science, Jeong, Songer, and Young Lee (2007) analysed forty sixth grade students' responses to the test of reasoning skills involved in the collection, organisation and interpretation of data contextualised in atmospheric science. They found that students' understanding of scientific evidence and reasoning skills regarding the data collection process were quite weak in several important aspects, including appreciating the importance of scientific evidence, identifying the relevant evidence, and properly interpreting examples and tables. By analysing argumentations within peer-led sessions in small groups in an undergraduate chemistry course, Ushiri, Moog, and Lewis (2013) found that although students could support their arguments with evidence and reasoning, their answers often lacked elaboration on reasoning and further validation on explanations by relating the argument with their prior knowledge. These studies show that students not only need help in conceptual understanding of scientific evidence, but more importantly they could benefit from explicit instruction of investigation practices that use scientific evidence.

Explicit instructional supports, including a well-designed learning environment and instructor's facilitation, are indispensable for guiding students to understand and practice using scientific evidence in order to develop their evidentiary reasoning skills (Manz, 2016). The importance of learning science through scientific investigation has been emphasised in the recent decade. Increasing encounters with scientific investigations in biology classrooms and laboratories could help students improve their understanding of core concepts (AAAS 2011), and grasp basic scientific competencies in experimentation (Pelaez et al. 2017; Pelaez, Gardner, and Anderson 2022) and cultivate biological literacy (AAAS 2011). The ability to understand and use scientific evidence as an essential component of scientific investigation is gaining focus as a foundation of undergraduate biology education (AAAS 2011; Laursen 2019). However, we do not yet know whether the trend towards engaging undergraduate students with scientific investigations yields gains similar to those from other instructional methods that have been reported to have a positive influence on learners' scientific reasoning ability (Gerber, Cavallo, and Marek 2001; Jensen and Lawson 2011; Wilson et al. 2010; Blumer and Beck 2019). The cultivation of competent thinking and evaluation of scientific evidence has become a crucial problem for educational research and instructors in higher education. Since scientific reasoning ability may not develop naturally among most students when exposed to traditional curriculum (Kuhn 2009), there is a need for including evidentiary reasoning as one of the learning objectives in science education.

Instructors also play a critical role in supporting students' learning through guiding students during scientific investigations. Instructors help students make connections between activities, science concepts and principles to support students' conceptual understanding (Puntambekar, Stylianou and Goldstein, 2007), they engage students in argumentation by justifying their claims, and they guide debating of alternative explanations (Osborne, Erduran and Simon, 2004; Tabak and Baumgartner, 2014). In this study, we focus on preparing instructors with explicit instructional methods to improve students' evidentiary reasoning and understanding of the role of relevant disciplinary knowledge to be called on during a scientific investigation.

HWE as a context for practising evidentiary reasoning

The HWE is a fundamental model for population genetics that was first demonstrated separately by two scientists, G. H. Hardy and Wilhelm Weinberg, in the early 20th century (Hardy 1908; Stern 1943; Weinberg 1908). According to this model, for a Mendelian trait that contains two alleles at one locus, one a dominant allele and another the recessive allele, the allele frequencies will remain constant across generations in a population if certain conditions are met (Figure in Supplemental Material 1). The conditions include no mutation in the gene, random mating between the individuals in the population, every individual produces the same number of offspring, no gene flow into or out of the population, infinite population size and no occurrences of natural selection. No evolution occurs in the population if it is in HWE, therefore understanding HWE is critical for grasping basic ideas and core concepts of evolution (Wise 2018b). Within biology, the HWE is a principle that is fundamental to the theory of evolution, which is firmly established as a model for helping learners to make sense of the importance of processes such as mutation, selection, and genetic change that would otherwise be meaningless for understanding the diversity of life (Dobzhansky 1973). For this reason, HWE is often taught as a fundamental introduction to population genetics in many introductory-level biology courses. By tracking changes in allele frequencies for a population over time, HWE helps students appreciate that evolution is not only the development of new species from existing ones, but that it can easily result from changes in the allele frequencies within a population over a long period of time. The topic of HWE in undergraduate biology education provides a foundation for the increasingly important discipline of evolutionary biology. By including the HWE, instructors also address the call to connect the use of mathematics with learning about biological phenomena (AAAS 2011; Schuchardt and Schunn 2016; Speth et al. 2010; Wise 2018b). Despite the importance of introducing HWE in undergraduate biology laboratory courses, several studies indicate that students show difficulty in understanding and applying HWE even when they have remembered and can use the HWE equation ($p^2 + 2pq + q^2 = 1$) for calculating problem solutions. Reports show that students have difficulty understanding the underlying biological phenomenon of relevance to HWE (Masel 2012; Smith and Baldwin 2015).

Our goal for this report is to show how the Conceptual Analysis of Disciplinary Evidence (CADE) framework, a novel conceptual framework on evidentiary reasoning, informed modifications of a laboratory investigation to focus on biological phenomena of relevance to the HWE in an undergraduate introductory biology laboratory course and how the implementation of scaffolding questions designed according to the CADE framework influenced the instructional questions about scientific evidence during the investigation. The CADE framework is introduced in the next section.

Theoretical framework

Scaffolding in science education

Scaffolding is a metaphor in education that refers to the process of teachers or knowledgeable peers using temporary supports to help learners complete tasks that may be beyond the learners' independent capacity (David, Bruner, and Ross 1976). Scaffolding is derived from Vygotsky's notion of the Zone of Proximal Development (ZPD) which refers to 'the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem-solving under adult guidance, or in collaboration with more capable peers' (Vygotsky 1978, 86). In the ill-structured problem-solving learning environment, with scaffolding from conceptual, metacognitive, procedural or strategic perspectives, learners can get support on what to consider, how to think during learning and tools or strategies for approaching the problem (Hannafin, Land, and Oliver 1999).

Prompts, hints and questions are useful scaffolding strategies. Instructional questions and prompts have been reported to have significant positive effects on students' performance on problem-solving and evidence-based argumentation, including representing the problem, identifying relevant information, gathering evidence to solve problems, generating hypotheses and linking evidence to claims (Belland, Glazewski, and Richardson 2008; Xun and Land 2003; van Joolingen and De Jong 1991). For example, in the context of information sciences and technology, Xun and Land (2003) provide examples of specific questions they used to prompt thinking about their work on a problem as students independently worked through and discussed ideas with peers. They found that students who received such questions and prompts were able to clearly represent the problem and identify the relevant information when solving the problem. However, the example questions used by Xun and Land (2003) do not specifically target the important role of evidence nor did those questions direct students towards evidentiary reasoning appropriate for a laboratory investigation. Therefore, as a first step to guide students' thinking and reasoning with and about scientific evidence during the process of scientific investigation, there is a need to clarify how scientists reason with and about scientific evidence throughout the entire research process.

The CADE framework

One useful framework for unpacking evidentiary reasoning is the CADE framework (Samarapungavan 2018). In contrast to other frameworks and studies on argumentation that focus on using evidence in making claims or explanations (e.g. Erduran and Pilar Jiménez-Aleixandre 2012), the CADE framework emphasises reasoning with and about evidence throughout the entire process of scientific investigation, including planning, generating and using scientific evidence. The CADE framework unpacks evidentiary reasoning into four relationships: the *Theory => Evidence relationship* refers to the practice of formulating testable hypotheses, explanations or rationale for an investigation; the *Evidence <=> Data relationship* refers to the practice of designing, executing, and analysing investigation models; the *Evidence => Theory relationship* refers to models of inference, argumentation and discussions about the uncertainty or sufficiency of conclusions, and *Social Dimensions* refer to the communication of evidence throughout the research process in a public sphere. These four relationships represent the entire stages of the scientific investigations, from identifying important unsolved problems, choosing theoretically important variables, collecting data to contribute as evidence to test hypotheses, drawing conclusions based on the evidence and communicating the research plans and results in the public or with other scientific researchers.

Within each of the four relationships, the CADE framework focuses on two essential components of reasoning with and about scientific evidence: *disciplinary knowledge* and *epistemic considerations*. Disciplinary knowledge provides the foundation for the investigation. The theories or assumptions that one person has will affect the person's decisions about what to choose as evidence, how to use the evidence, and what can be drawn as conclusions from the evidence. Epistemic considerations throughout all four relationships of evidentiary reasoning relate to critical thinking about the scope, quality and limitations of the theories, evidence and conclusions. When practising a scientific investigation, investigators need to not only grasp the disciplinary knowledge that is necessary to conduct the investigation, like the theories and models that are underlying the investigation, the variables that are relevant to the research question, but also how to evaluate the quality of the scientific evidence, including the reliability of the evidence, as well as the precision and accuracy of the techniques and equipment used. Thus, the disciplinary and epistemological aspects of evidentiary reasoning are interrelated when conducting evidentiary reasoning. From the educational perspective, instructors need to not only introduce students to the disciplinary knowledge that is necessary to conduct the investigation, but they also need to implement curriculum and instructional methods to support students' epistemic considerations for preparing students' critical thinking needed for both scientific learning and daily life (Barzilai and Chinn 2018; Chinn, Barzilai, and Golan Duncan 2020). In each relationship and under each component, the CADE framework poses several questions regarding important features of scientific evidence. It could benefit instructors in developing questions and prompts to scaffold students' evidentiary reasoning during scientific investigations. Although the CADE framework is built on examples from biology, the proposed questions, especially questions of relevance to epistemic considerations, are very general and are intended to be applicable in many different contexts and disciplines.

In summary, the CADE framework is intended to provide a guide for both instructors and learners to understand the construct of scientific evidence and evidentiary reasoning. Since the study context is an introductory biology laboratory course, we only focus on the first three CADE framework relationships in this article, as these represent what happened during the HWE laboratory investigation. Here we report on CADE implementation with a three-part study to address these questions: What are the features of a lesson plan for an HWE investigation that is informed by CADE? How does an instructor's implementation of scaffolding for an HWE investigation compare before and after implementation of CADE? And how does our CADE-informed HWE laboratory investigation compare to other published activities for engaging students with HWE investigations?

Study methods and design

Study context: A population investigation module with zebrafish

This study was conducted in an introductory biology laboratory course at an American research-intensive institution in the Midwest during three continuous semesters with 2018 Spring semester as the baseline course and 2018 Fall and 2019 Spring semesters as two iterations. The design of the laboratory course was intended to engage students who were first year biology majors in practising hypothesis testing repeatedly as they progressed through a series of research modules. One 50-minute pre-lab lecture each week was delivered to about 200 students by the professor to introduce concepts and techniques that would be used in their 3-hour lab meeting where about 18 students came together to conduct investigations as members of one of six research teams per lab. Table 1 summarises student demographic information from 2018 Spring, 2018 Fall and 2019 Spring semesters. Data for this study was from the Spring semesters.

The zebrafish module was the second of four research modules in the course, and it lasted four weeks; the course design is presented in the Supplemental Material 2. During this module, students worked in groups of three to raise zebrafish (*Danio rerio*) and to collect data on their fish populations to test HWE with assigned fish tanks. In brief, some wild-type fish had zebra stripes, the dominant trait. Others had a recessive phenotype (spotted or without stripes). By analysing the changing proportions of zebrafish phenotypes, the students might conclude if evolution had occurred in the population in their tank. Each lab section consisted of one graduate teaching

Table 1. Student demographics in the biology laboratory course for each semester of this study.

Demographic variables	2018 Spring	2018 Fall	2019 Spring
Female	65%	61%	73%
Male	35%	39%	27%
White	65%	66%	66%
Black	4%	4%	3%
Hispanic	9%	6%	5%
Asian	11%	16%	15%
Multiracial	5%	5%	5%
International	5%	4%	4%
Unknown	1%	0	1%

assistant and 3–4 undergraduate teaching interns as teaching staff who guided the 18 undergraduate students. Shim and Pelaez (2022) describe more detail about the laboratory context for this study as well as examples of guidelines from the 4-week Zebrafish Lab Module. In the third week of the zebrafish module, students worked on a guided inquiry HWE activity in a lab meeting that lasted three hours. This study is based on data from that HWE lesson, so we will focus only on the third week of the module (Supplemental Material 2). During this lesson, the lab instructor used whole-class lab discussions to prompt students to consider some evidence of relevance to an investigation of HWE.

Data collection and analysis methods

Modifications to the lesson for an HWE investigation were informed by CADE

Volunteer teaching staff members were invited to collaborate with the first and the last authors to improve teaching strategies according to the CADE framework by modifying the lesson for the HWE laboratory investigation between semesters. The HWE lessons were modified twice from the baseline course (Spring 2018 semester) following two iterations informed by CADE (2018 Fall and 2019 Spring semesters). The final lesson in the 2019 Spring semester was coded with the three relationships and the two knowledge categories of the CADE framework. The scaffolding questions designed at the beginning of the first iteration were modified and applied for the final iteration by the authors in collaboration with the participant lab instructors.

Participant instructors who contributed to course improvements were paid for their participation. So that the laboratory discussions could be recorded, the laboratory instructor and student participants agreed to participate according to a protocol that was reviewed and approved by the Institutional Review Board (IRB#1702018760251). Scaffolding questions were modified for the final iteration based on observations of classroom discussions during the baseline and first iteration.

Instruction comparisons before and after CADE implementation

A participant instructor who was an undergraduate teaching intern participated in all three semesters. She earned a bachelor's degree in biochemistry in 2019 while completing minors in both Statistics and Biotechnology with a perfect 4.0 Grade Point Average. During her undergraduate career, she also participated as science research investigator in several research projects. She had been teaching the same laboratory course for one semester before volunteering to participate in our HWE laboratory improvements. Her excellent academic performance and the experience of teaching and research were aligned with her quick ability to understand and apply the CADE framework while helping to devise scaffolding prompts for a modified HWE laboratory lesson for each iteration. In the results section, her laboratory instructions provide examples to illustrate the changes made to instructional questions and prompts based on the CADE framework for guiding students' thinking and reasoning about scientific evidence during the HWE laboratory investigation.

The changes in this participant instructor's scaffolds were tracked by videotaping her laboratory courses to compare the baseline course in spring semester 2018 to the modified implementation informed by the CADE framework in the fall 2018 and spring 2019 semesters. For each course, the pre- and post- the lab activity discussions were videotaped (almost 1 hour each), during which the instructor implemented questions and prompts to lead the whole class discussion. The videotaped sections were transcribed using Trint.com, which is an online transcription tool. The transcripts were proofread by the first and the last authors. The CADE framework was applied as an analysis framework to indicate what aspects of evidentiary reasoning had been targeted for discussion during the HWE laboratory investigation and to identify what was still missing. Words and sentences from the statements, prompts and questions that instructors used were coded according to three relationships and the two components of the CADE framework. The description of codes and examples from the course data are presented in Table 2. Two independent coders (the first and the last authors) worked on 25% of the transcripts to refine the code book until agreement was reached on these codes. Then one of the coders (the first author) coded the rest of the transcripts.

Table 2. Code descriptions and examples.

Code	Description	Example*
Theory => Evidence relationships, Disciplinary Knowledge	Disciplinary knowledge about formulating or testing hypotheses, explanations or arguments	(Baseline course) *'So if natural selection is occurring, how will that affect how the offspring compare to the parent generation?'* (The HWE dog investigation) *'So, do you guys have any hypotheses on whether or not you expect your allele frequencies to be similar or different than in New York?'*
Theory => Evidence relationships, Epistemic Considerations	Ideas about the nature, quality and scope of evidence in formulating or testing hypotheses, explanations or arguments	(The HWE dog investigation) '*Can you guys think of any alternative methods that if maybe the time or money wasn't an issue for us to kind of solve this problem*?'
Evidence <=> Data relationships, Disciplinary knowledge	Disciplinary knowledge about designing, executing, analysing investigation models	(Baseline course) '*So, hearing that you have to do the expected and observed frequencies, what do you guys think you'll have to do later that we talked about last lecture?*' (The HWE dog investigation) '*If you have the data of these different genotypes, how do you convert those to allele frequencies?*'
Evidence <=> Data relationships, epistemic consideration	Ideas about the nature, quality and scope of evidence in designing, executing, analysing investigation models	(The HWE dog investigation) '*How might you want to increase your confidence in this data?*'
Evidence => Theory relationships, Disciplinary knowledge	Disciplinary knowledge about drawing conclusions from the evidence	(The HWE dog investigation) 'W*hat does this data tell us in the context of our model?*'
Evidence => Theory relationships, Epistemic consideration	Ideas about the nature, quality and scope of evidence when drawing conclusions	(The HWE dog investigation) '*So any other big limitations that you guys might want to try to remediate?*'

*Quotes are the words of the instructor with only slight editing for grammar and removing filler words that do not change the meaning of the quote (like, kinda) to make the quotes easier to read.

Comparing published HWE activities

To compare our modified approach to what others are doing to engage students with HWE investigations, published HWE instructional activities were examined. Searching was performed in Google Scholar to find recently published activities or investigations for teaching Hardy-Weinberg Equilibrium. The searches were conducted using two sets of key words: 'teaching Hardy Weinberg equilibrium' and 'Hardy Weinberg equilibrium instruction'. Articles published in peer-reviewed journals from 2000 to 2019 were chosen. The activities were compared according to several instructional features, including the learning objectives, the data sources used for solving HWE problems, whether the activity provides detailed disciplinary knowledge of relevance to the context of the HWE investigation in order to consider the disciplinary aspects of evidence according to the CADE, whether the activity encourages students to ask a research question that they can address with their evidence, and whether the activity includes scaffolding questions to alert students to consider disciplinary knowledge of relevance as well as general evidence considerations for the instructor to guide students to use evidence better in the HWE activities.

Results

Reasoning about evidence from dog populations

Here we present a novel HWE laboratory investigation of dog populations to give undergraduate biology students a guided project with an investigation of the HWE (Box 1), a practical guide with scaffolding questions (Table 3) for instructors to implement in their courses, and an assessment which was designed to measure students' competence for evidentiary reasoning for HWE on the final exam at the end of the semester (Box 2).

Box. 1. A dog population investigation of HWE.

Consider two phenotypes for which the genetic mechanisms in dogs are known: the black colour trait is related to variation in *Tyrosinase Related Protein 1* (*TYRP1*) alleles and a trait for the type of dog facial fur is related to *R-spondin 2* (*RSPO2*). The combination of two mutations of *TYRP1*, symbolised by *b/b*, alters eumelanin from black to brown and gives dog fur and skin the 'brown' phenotype, which shows up as no black colour on nose, paw pads, eye rims or lips (Figure 3). The 'black' phenotype is due to the dominant allele and can be symbolised by *B/_* (Dreger et al. 2019; Jancuskova, Langevin, and Pekova 2018; Schmutz and Berryere 2007; Sheila M Schmutz, Berryere, and Goldfinch 2002). Scruffy fur phenotype on a dog's face is related to Roof Plate-Specific Spondin-2 (*RSPO2*), which encodes an extracellular matrix signalling pathway protein. The *RSPO2* allele for the scruffy trait is due to the dominant allele, illustrated by *F/_*, in contrast to smooth facial fur on dogs that are homozygous recessive for this trait, illustrated by *f/f* (Figure 4) (Cadieu et al. 2009; Dreger et al. 2019). A website that lists adoptable animals across most of North America is http://PetFinder.com. Using Petfinder.com, a scientist of canine genetics found that in New York City the gene pool for adoptable dogs is mostly smooth faced (a recessive trait only seen in dogs that inherit two copies of the *RSPO2* gene for smooth face). But many dogs carry the *TYRP1* gene mutation that causes some dogs to be missing black pigmentation. On 1 February 2019 in New York City, the counts of *RSPO2* and *TYRP1* phenotypes were as follows:

Smooth Black	Smooth Brown	Scruffy Black	Scruffy Brown
f/f B/_	*f/f b/b*	*F/_ B/_*	*F/_ b/b*
80	22	15	1

Assuming Hardy-Weinberg Equilibrium:

- q^2 = frequency of the recessive phenotype
- q = frequency of the recessive allele
- p = frequency of the dominant allele → p = 1 – q
- For *TYRP1* brown:
  - q = *b* = brown = 0.442
  - p = *B* = black = 0.558
- For *RSPO2* scruffy:
  - q = *f* = smooth face = 0.9302
  - p = *F* = scruffy face = 0.070

What research question can be investigated with this and other data from http://PetFinder.com?

Table 3. Practical guidelines for discussing dog population investigations with scaffolding questions informed by the CADE framework.

Stage	Align with the CADE framework	Instructional practices	Scaffolding questions
Before the investigation: Reasoning with and about the HWE model to prepare for the investigation.	**Theory => Evidence relationships** Model Articulation: Formulate/test hypotheses or pose explanations and a rationale for the investigation	A pre-laboratory lecture helps students understand the Hardy-Weinberg Equilibrium (HWE) Model and the canine genetics related to the dog population investigation. The instructor guides the student investigators to apply the HWE model to a dog population using the dog data from petFinder.com with New York City as a worked example to benefit investigators who lack experience with applying maths to biological problems, who lack motivation to solve problems, or who lack contextualised knowledge about how to gather and organise data to apply to the Hardy-Weinberg Equilibrium model. During the lab session, the instructor guides students to revisit disciplinary knowledge about the HWE and canine genetics. With guidance the class poses a research question they will answer. For example, students may want to find out if there has been a change in allele frequencies of *TYRP1* and/or *RSPO2* for New York City dogs over time or they might be curious about whether dogs in their hometown differ from the New York dogs. The scaffolding questions highlight both disciplinary knowledge and epistemic considerations to think and reason about the HWE model and their rationale for their research question to generate testable hypotheses for a strong investigation.	**Disciplinary knowledge** ● Why do biologists care about the HWE model? Why use it for an investigation? (What are the key domain phenomena?) ● How do the assumptions for Mendelian genetics differ from those for populations with HWE? ● What would a biologist think counts as evidence and what sort of data is collected according to HWE? (What variables are relevant?) ● Pose a research question to be answered about a population and state the null hypothesis. (What are the important unsolved problems?) **Epistemic considerations** ● Considering the various models (HWE or Mendelian genetics), which is most appropriate to address your research question and why? (Were alternatives evaluated?) ● What limitations are associated with the model? (Is the articulated model complete, specific, and internally consistent?)

(*Continued*)

Table 3. (Continued).

Stage	Align with the CADE framework	Instructional practices	Scaffolding questions
During the investigation: Sample the dog phenotype data from the Petfinder.com and testing of Hardy-Weinberg Equilibrium or comparing two populations with Chi square test.	**Evidence <=> Data relationships** Designing, executing, analysing evidence from investigations	The instructor leads a discussion for student investigators to develop a non-biased sampling strategy. Using Petfinder.com the student investigators apply and refine a sampling strategy to collect data, they design a data table to compile all the findings and they discuss how to use the data as evidence to answer their research question. The investigators ca apply chi-square test with the 'expected versus observed' counts data (HWE) or to compare two conditions (dogs in California versus New York, for example). With scaffolding questions from both disciplinary knowledge and epistemic considerations, the instructor can prompt students to think and reason with the data including the collection and analysis procedures to consider the scope and quality of the evidence they generate.	**Disciplinary knowledge** ● How will the data be used as evidence for testing your model? ● What statistical analysis models are used to organise/compare data to see whether or not a population is in Hardy-Weinberg Equilibrium? (What analysis models are used to organise/compare data (e.g. graphs, statistical models)?) ● Given the HWE assumptions, what are the known sources of error and how will they be accounted for? (What are known sources of error and how will they be accounted for?) ● Why did you set this specific alpha-level for the statistical test? Would other levels be appropriate? (What data reporting standards apply? E.g. attrition, error rates, outliers?) ● What analysis models are used to compare data? (What analysis models are used to organise/compare data (e.g. graphs, statistical models)? **Epistemic considerations** ● Are the technical data collection procedures adequate? (Is technical precision, power, sensitivity, reliability, data collection procedures adequate?) ● Considering the sampling procedures, are data sampled in an unbiased way that is representative of the population? (Are data sampled in an unbiased way and representative of the data population?) ● What do you think of the quality of your data for answering the research question? (Are data robust?) ● What additional evidence would give more confidence? (Are data reports fair/complete?)

(*Continued*)

Table 3. (Continued).

Stage	Align with the CADE framework	Instructional practices	Scaffolding questions
After the investigation: Interpret the evidence and draw the conclusion.	**Evidence => Theory relationships** Models of argument: Sufficiency of conclusions	Student investigators describe the populations they are comparing assuming each is in HWE, and they consider limitations of this assumption in drawing conclusion from the evidence they generated from their investigation when the instructor leads a discussion about the interpretation of the results. The scaffolding questions consider both disciplinary knowledge and epistemic considerations to prompt students' thinking and reasoning about the strength of their evidence interpretation and the sufficiency of their conclusion or claims.	**Disciplinary knowledge** ● What has been learned about the population from the evidence? (What has been learned from the evidence?) ● When performing the proposed hypothesis test, what does it mean if you fail to reject your null hypothesis? (What has been learned from the evidence?) ● Are the findings consistent with the idea that the population evolved? (Are findings/conclusions consistent with the totality of disciplinary evidence available?) ● What would a biologist think could be a likely explanation for the findings? (Are there alternate interpretations that are compatible with the evidence?) **Epistemic considerations** ● How confident are you in this interpretation? (Are conclusions internally consistent?) ● How confident are you that H_0 should have been rejected/not rejected? (Are conclusions internally consistent?) ● Are there any alternative interpretations to explain the findings? (Are alternate interpretations evaluated?) ● Do the findings fit with other reports from published studies? (Is fit with other studies considered?)

Note. Questions underlined are original questions from the CADE framework.

Box 2. A dog population investigation assessment.

According to 'Dogs', a series on Netflix, *Territorio de Zaguates* **is a free-range shelter run by people who are dedicated to saving dogs abandoned in Costa Rica. The phenotypic traits in the shelter were well analysed. The head vet of the shelter stated that these dogs are unique at Costa Rica**.

For comparison, a scientist of canine genetics, found that the gene pool for adoptable dogs is mostly smooth faced (a recessive trait only seen in those that inherit two copies of the *RSPO2* **gene for smooth face), but many dogs carry the** *TYRP1* **gene mutation that causes some dogs to be missing black pigmentation. Using a website that lists adoptable animals across most of North America is http://PetFinder.com these counts of** *RSPO2* **and** *TYRP1* **phenotypes were found on 1 February 2019, in New York City:**

Smooth Black	*Smooth Brown*	*Scruffy Black*	*Scruffy Brown*
f/f B/_	*f/f b/b*	*F/_ B/_*	*F/_ b/b*
80	22	15	1

Questions

a) What would you observe about the dogs in the Netflix show, *Territorio de Zaguates*. to find out if those dogs are really different from dogs that are up for adoption in New York City? Is there evidence that the gene pool for these dogs is different in Costa Rica? Explain an approach to doing research on the dogs. What should the investigators be looking at to determine their uniqueness? Why would these observations be interesting and/or important?

b) The vet has described the phenotypes of 1000 dogs at the *Territorio de Zaguates* free-range shelter. What types and how many dogs with each phenotype would you expect to find in this data if the population does NOT differ in terms of the allele frequency for the traits this scientist reported for adoptable pets in New York City? Assume both populations are in Hardy-Weinberg Equilibrium.

c) Given what you know about Mendel and the Hardy Weinberg Equilibrium principle, what questions do you have and what additional evidence would help you figure out what is happening with the abandoned dogs in Costa Rica?

d) What else might you need to know to improve the quality and the accuracy of the vet's claim that the abandoned dogs in Costa Rica are unique?

Before the laboratory session, a 50-minute pre-lab lecture was delivered by a guest professor to review HWE and to introduce colour variation in dogs associated with *Tyrosinase Related Protein 1* (*TYRP1*) alleles (Figure 1) and variation in the type of dog facial fur related to *R-spondin 2* (*RSPO2*) alleles (Figure 2). For this HWE laboratory investigation, Petfinder.com was used as a data source, which is a free online searchable database for adoptable pets in certain areas by the postal zip code. The scaffolding questions

Figure 1. Comparison of the 'black' and 'brown' colour phenotypes in dog populations caused by variation in *Tyrosinase Related Protein 1* (*TYRP1*) alleles. Dogs with the 'brown' colour trait lack any black pigmentation as illustrated in photos marked with a green check mark to illustrate the 'brown' colour phenotype (symbolised by *b/b*). This lack of black pigmentation is caused by the combination of two mutations of *TYRP1* and shows up as no black colour on nose, paw pads, eye rims or lips. The photos marked with red cross illustrate the 'black' colour phenotype (symbolised by *B/_*). The two photos in the lower right corner demonstrate how a genetically black (*B/_*) nose can fade over time. This should not be mistaken for a brown nose (*b/b*) as the dog remains genetically 'black' despite the fading.

Figure 2. Comparison of the scruffy and smooth face phenotypes in dog populations related to variation in *Roof Plate-Specific Spondin-2* (*RSPO2*) alleles. The dogs with scruffy face phenotypes (symbolised by *F/_*) are in photos marked with green check marks. *R-spondin 2* (*RSPO2*) encodes an extracellular matrix signalling pathway protein. The dog photo marked with a red cross shows the recessive smooth face phenotype (symbolised by *f/f*). The two photos in the lower right are of the same dog, with a groomed face and with an ungroomed face. These are shown as an example of how the appearance of scruffy face can be altered with grooming but, of course, this does not alter the alleles.

present in the lesson (Column 4 of Table 3) involve figuring out an investigation to determine how the allele frequencies of a Mendelian trait in an observed population differ from a comparison population in a different place or the same population at a different time, using the phenotype frequency to predict allele frequencies as evidence for evolution. Instructors can apply these scaffolding questions at various parts of the lesson to prompt deeper thinking and reasoning with and about scientific evidence while guiding students through the investigation. For example, before the investigation, the instructor can use the question '*What important unsolved problem could be addressed with HWE in your investigation?*' (Theory = > Evidence relationships, Disciplinary knowledge) to prompt discussion of a research question that the student teams could investigate. Note that the *Disciplinary knowledge* component in Table 3 recognises specific disciplinary considerations to inform the investigation, such as HWE, Mendelian genetics, and relevant statistical models. In contrast, the *Epistemic considerations* in Table 3 raise questions about the nature, scope and quality of scientific evidence to support epistemic reasoning about evidence with more generalised strategies.

The assessment in Box 2 presents a problem with open-ended questions that include and go beyond simple calculation with HWE equations. The assessment questions are designed to reveal students' thinking about the entire investigation process in terms of both disciplinary knowledge and epistemic considerations. For example, the question '*Given what you know about Mendel and the Hardy Weinberg Equilibrium principle, what questions do you have*?' is aligned with the *Theory => Evidence relationship*. Student reasoning about *Evidence <=> Data relationship* is revealed when they consider '*What additional evidence would help you figure out what is happening with the abandoned dogs in Costa Rica*?', and the question '*What else might you need to know to improve the quality and the accuracy of the vet's claim that the abandoned dogs in Costa Rica are unique?*' is related to the *Evidence =>Theory relationship*. The assessment was designed to reveal students' understanding of the HWE model, their competence with hypothesis testing and their reasoning about the nature, scope, and quality of the data that underpin their use of scientific evidence.

Comparison of scaffolds for HWE investigation laboratory instruction before and after the CADE implementation

Changes from the participant instructor's lab instruction between the baseline course and the course that implemented the dog population investigation are presented to show how the CADE framework influenced the instructional questions and prompts about evidentiary reasoning when the participant instructor guided discussion of the HWE lab investigation. The counts of different types of questions used by the instructor before and after implementing the CADE framework are summarised and presented in Table 4.

Scaffolds for laboratory discussion during the baseline course

During the baseline HWE activity, students were given phenotype numbers to calculate both allele frequencies and genotype frequencies (Figure 3). Alignment with the CADE framework revealed that questions and prompts used by the participant instructor during the baseline activity only covered two relationships in the CADE framework, which are the *Theory => Evidence* and *Evidence <=> Data relationships* (Table 4), and even in these two relationships, some important components were missing. Most of the questions in the *Theory = > Evidence relationships* were intended to refresh students' memory about the HWE assumptions (5 out of 8 in Theory => Evidence relationships, Disciplinary knowledge in Table 4). For instance, the instructor said, ' … *the population has to be very large. Can anybody give any predictions about why that might be?*' (Theory => Evidence relationships, Disciplinary knowledge). Students then discussed that allele frequencies in the large population would unlikely be affected by random events since one individual makes up a lot less of a large population. One student answered ' … *in very large populations, one individual makes up a lot less of a population*'. There were no questions to prompt discussion about a key domain phenomenon, which is that the appropriate model is HWE and not just Mendel's Laws applied to population genetics. No questions were implemented to guide students to compare the two genetic models that they had learned, or to discuss how Mendel's Laws or HWE fit different types of investigations. This led to the result that some students failed to select the proper model to address a population genetics study. Based on their responses to the assessment (Box 2) on the final written exam for the spring semester 2018, over one third of the students applied Mendel's Laws and used Punnett squares to calculate the genotype frequencies in solving a population genetics problem, instead of considering p or q as allele frequencies (unpublished data).

Since the activity in the baseline course focused on applying HWE in terms of an equation with a given set of numbers, most of the guiding questions in the *Evidence <=> Data relationship* were simply about how to calculate the allele or genotype frequencies with given numbers (11 out of 12 in Evidence <=> Data relationships, Disciplinary knowledge in Table 4). For instance, the instructor asked ' … *so if you have 40% aa and then 60% of AA,*

Table 4. Counted differences in the lab instruction between the baseline course and the HWE dog population investigation.

	Baseline		Dog population investigation	
Constructs and Practices	Disciplinary knowledge	Epistemic considerations	Disciplinary knowledge	Epistemic considerations
Theory => Evidence relationships	8	0	15	2
Evidence <=> Data relationships	12	0	16	7
Evidence => Theory relationships	0	0	13	9
Total	20	0	44	18

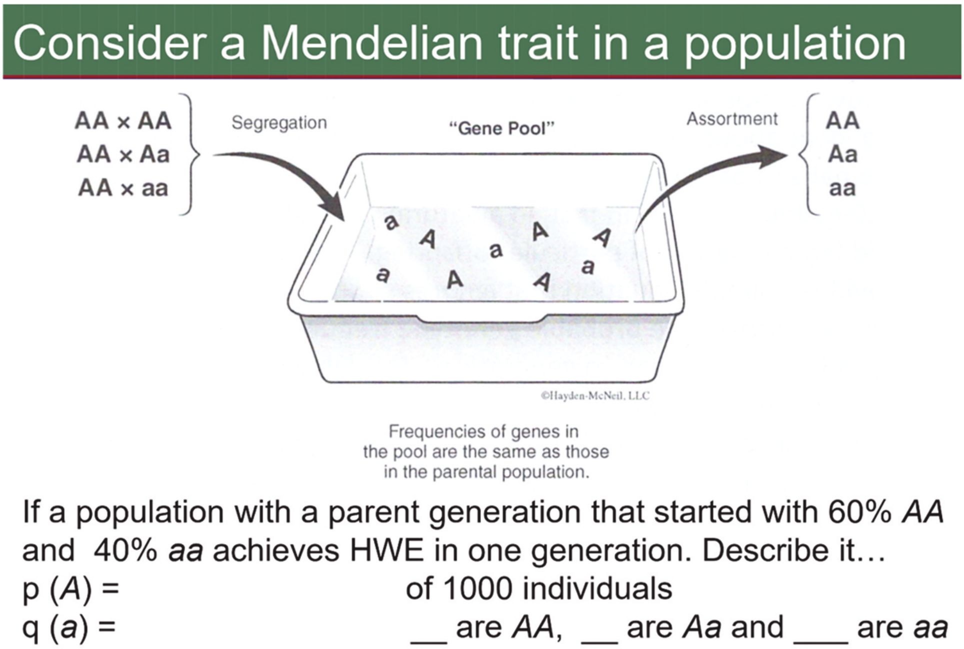

Figure 3. Instructional diagram for the Hardy-Weinberg equilibrium (HWE) laboratory activity in the baseline course. The baseline HWE laboratory activity illustrated allele frequencies in the gene pool and focused on helping students calculate with the HWE equations using numbers of genotype frequencies of a trait that were given to students.

what would the proportion of just the A allele be in the population?' (Evidence <=> Data relationships, Disciplinary knowledge). Students then applied the calculation of HWE equations with the given numbers and gave their answer as '*0.6 for the dominant A allele frequency*'. In the baseline lab activity, there were no opportunities for students to discuss and critique the sampling methods, or the accuracy of the methods and techniques used to gather the data, which are essential for scientific investigations.

Most of the questions in the baseline course were intended to prompt students' thinking about disciplinary knowledge related to the HWE activities. There were no guiding questions to prompt students' epistemic considerations about the quality or reliability of evidence used in the activity (Table 4). This illustrates the need to modify the lesson and design scaffolding questions to prompt students' epistemic considerations during their HWE laboratory investigation to encourage students to think about the nature, scope and quality of the evidence used during the investigation.

Scaffolds for laboratory discussion during the HWE dog investigation

After implementing the dog population investigation and incorporating scaffolding questions into the HWE lesson, several major changes were observed in the laboratory instruction. First, there were more questions and prompts for evidentiary reasoning corresponding to the first three relationships in the CADE framework (62 vs 20 in Table 4). This indicates that the instructor intended to encourage students to think and discuss multiple aspects of scientific evidence during the lab investigation. These questions covered important components for evidentiary reasoning that were missing from the baseline laboratory instruction. For example, in order to help students compare Mendel's Laws to the HWE and to choose the proper model for the investigation, the instructor said, '*We have learned about two* *genetic models – Mendelian genetics and population genetics with Hardy Weinberg Equilibrium. What are the assumptions for both of them?* (Theory => Evidence relationships, Disciplinary knowledge). *Can you use either model? Which one is more appropriate and why?*' (Theory => Evidence relationships, Epistemic considerations). These questions led the students to compare the two genetics models they had learned and to think about the relevant model and mechanisms underlying changes of allele frequencies in a dog population. Students replied that the HWE is the model they were going to apply for the population investigation. They recalled the Mendelian inheritance session they've learned before and identified the difference between the two-investigation contexts as that ' ... *for the Mendelian inheritance, it has to come from one set of parents. But for the HWE model, we are using here with the pet populations, means there are numerous sets of parents that could all be interbreeding*'.

The features of the dog investigation encouraged students to participate in the entire hypothesis testing process, thus providing a context for the instructor to implement the CADE framework with scaffolding questions to inspire students' evidentiary reasoning regarding in identifying a research question, generating hypotheses, collecting and interpreting data as evidence and drawing conclusions. For example, before the investigation, the instructor asked ' ... *so, what research question are you interested in? Do you have any hypotheses on whether or not you expect your allele frequencies to be similar or different than in New York?*' (Theory => Evidence relationships, Disciplinary knowledge). One student mentioned that his hometown in a rural area of North Carolina is prone to flooding. He wondered whether the fur type of the dog which related to *R-spondin 2* (*RSPO2*) in his hometown would differ from a bigger city, like New York City, that is not rural or so prone to flooding. The class was interested in this research question and decided to investigate whether there were statistically significant differences in allele frequencies of smooth face fur (*f*/*f*) between a rural location in the North Carolina flood zone and New York City. During the investigation, the instructor led discussions about evaluating the sampling methods and details of the statistical methods applied to analyse the data. This let the whole class discuss how to process an unbiased sampling method; for example, how many dogs each of them would count and how they could avoid counting the same dog twice on Petfinder.com as evidence. After processing the data to generate evidence, the instructor asked, '*So what conclusions can you draw?*' and then '*So, we've rejected our null hypothesis, now what do we say about our model and our data?*' (Evidence => Theory relationships, Disciplinary knowledge). The whole class used evidence to draw their conclusion that there was no significant statistical difference between the dog populations in a rural location in the North Carolina and New York City. They then discussed potential reasons and alternative interpretations for this conclusion.

Another change was that with the implementation of scaffolding questions of epistemic considerations, the instructor started to encourage the students to talk about the limitations and to evaluate the theory and evidence of relevance to the dog population investigation. The instructor guided students to think about limitations of the HWE by asking ' ... *but any other limitations that you might think of for this sort of model ... ?*' (Theory => Evidence relationships, Epistemic considerations). The class discussed the scope and quality of the evidence collected by students and used during the investigation when the instructor asked, '*How much confidence do you guys have in this evidence for allele frequencies? So, does anybody have any ways and ideas of how you might want to increase your confidence in this data?*' (Evidence <=> Data relationships, Epistemic considerations). Students criticised the quality of the data they collected. Some students mentioned the data collection process that '*we probably counted the same dog multiple times when collecting the data*'. Some mentioned the sample size that '*we only did forty out of the entire state*'. These changes support the idea that implementation of the CADE framework can help an instructor use scaffolding questions that cover epistemological aspects of evidentiary reasoning to guide students' thinking and reasoning with scientific evidence in a way that is integrated into the disciplinary context of their HWE laboratory investigation.

Activities for teaching HWE in undergraduate biology courses fail to address some CADE components

Table 5 presents a comparison of our dog population laboratory investigation (bottom row) with several published classroom or laboratory activities designed for teaching HWE. Although varied in learning objectives, the activities listed in Table 5 were intended to shift the teaching of HWE from a traditional focus on calculating with HWE equations towards engaging students to understand the biological phenomena underlying the HWE by involving students with an investigation. For example, in some hands-on activities, students participated in modelling HWE using coloured beads or papers and calculated frequencies of alleles, genotypes and phenotypes to simulate changes in allele frequencies within a population (Bonner et al. 2019; Brewer and Gardner 2013; Hoffman et al. 2016; Smith 2017). This engages students with analysis of the data but gives no opportunity for students to discuss which model (HWE or Mendelian genetics) is appropriate and why. A cat population activity (Christensen 2000) that was useful to inform our dog HWE activity is based on data that students collect as evidence to test HWE, as was suggested previously by Kinnear (1986) for teaching Mendelian genetics. However, the cat as a research subject raises questions about the detailed genetic disciplinary knowledge that determines the phenotype of cats (Kinnear 1986; Christensen 2000). Other than the dog population investigation, the published activities listed in Table 5 for teaching HWE do not provide opportunities for students to think about the quality of scientific evidence, a practice they would need to understand and use for an authentic scientific investigation. However, even though several of these approaches report hypothesis testing as one of the learning outcomes, they lack detail about the disciplinary knowledge needed for evidentiary reasoning of relevance to formulating or testing hypotheses. Even published activities that have students compare simulation results with authentic research data are simplified. Students are not getting opportunities to engage in experimental design with these activities in ways that involve discussion of sampling procedures and research methods, as well as learning about evidence in terms of the nature, scope and quality of evidence to develop their evidentiary reasoning ability.

The dog population investigation was designed to engage students in designing and planning a research study, developing methods for sampling, independently generating authentic data as evidence, and drawing conclusions from the analysis. In addition to involving students in practicing hypothesis testing, more importantly, the task is intended to inspire students to think and reason with and about scientific evidence with scaffolding questions focused on both epistemic and disciplinary aspects before, during, and after the investigation. In contrast to other studies in Table 5, students could apply their knowledge of molecular biology to data they collect from Petfinder.com along with data a scientist gathered about dogs in New York City to answer a research question about whether there is a statistical difference in the allele frequencies for *TYRP1* or *RSPO2* between two cities, for example New York City and their hometown. As an alternative, the students could have decided whether there is difference in allele frequencies for *TYRP1* or *RSPO2* over time, for example before and after the Covid-19 pandemic in New York City.

Table 5. Comparison of activities published for teaching and learning HWE from a laboratory investigation.

Activity	Author(s)	Year	Course level	Data sources used for solving HWE problems	Details disciplinary knowledge (Yes/ No)	Raises a research question (Yes/No)	Provides scaffolding questions (Yes/No)	Activity Objectives and Highlights
Clam spawning & red tide: Helping students learn the Hardy-Weinberg Equilibrium	Bonner et al.	2019	Introductory- and senior-level undergraduate biology courses	(1) Simulation with coloured beads to represent different alleles. (2) Authentic research data published in a journal	Y	N	N	In ‘hands-on’ activities, students use beads for simulation and calculation. Simulation results are validated with authentic research data to provide a meaningful learning experience.
How to teach the Hardy-Weinberg Principle using engaging, non-trivial evolutionary scenarios	Wise	2018a	Introductory undergraduate biology course	Fabricated numbers	Y	N	N	Provides a four-step guide for teaching HWE, together with six evolutionary scenarios in which students apply the HWE. The complex scenarios are intended to benefit student’s quantitative reasoning and higher-order cognitive skills.
Teaching the Hardy-Weinberg Equilibrium: A 5E lesson plan	Smith	2017	High school biology, AP biology and introductory undergraduate biology course, with students who had some prior modelling experience	(1) Simulation with coloured paper clips to represent different alleles. (2) Authentic research data published in a journal	Y	N	N	The focus is on student’s understanding of the ‘big ideas’ about the HWE and evolution. The lesson provides students with disciplinary knowledge of the HWE, a research problem, and authentic data from published papers.
Development and assessment of modules to integrate quantitative skills in introductory biology courses	Hoffman et al.	2016	Introductory undergraduate biology course	Simulation with red and white kidney beans to represent the genotypes in a population of rabbits	N	N	Y	Learning goals include demonstrating quantitative numeracy, interpreting data sets, communicating the interpretations with visual tools, making statistical inferences from data sets, and explaining how evolutionary mechanisms contribute to change in gene frequencies in populations. Using a hypothetical recessive trait, no disciplinary knowledge is provided about the target trait or gene.

(*Continued*)

Table 5. (Continued).

Activity	Author(s)	Year	Course level	Data sources used for solving HWE problems	Details disciplinary knowledge (Yes/ No)	Raises a research question (Yes/No)	Provides scaffolding questions (Yes/No)	Activity Objectives and Highlights
Teaching evolution through the Hardy-Weinberg Principle: A real-time, active-learning exercise using classroom response devices	Brewer and Gardner	2013	Nonmajors undergraduate biology course	Simulation with coloured papers to represent different alleles	N	N	Y	The focus is on improving student's conceptual understanding especially about the violations of HWE. No disciplinary context knowledge is provided about the target trait or gene.
Cats as an aid to teaching genetics	Christensen	2000	One-semester course consists of sophomores and juniors who have biology background	Students use a handout to determine the genotype at seven unlinked loci of at least one cat. Instructor collected students' data and used it to illustrate HWE.	N	N	N	This activity reinforces students' conceptual understanding, allows the students to look at population genetics in a very positive light and provides concrete examples of some misunderstood principles. However, using cat as a research subject raised questions about the detailed genetic disciplinary knowledge that determines the phenotype of cats. It is challenging to connect single genes to the observed phenotypes in the cat population.
The HWE laboratory investigation on dog populations	Authors of this study	2022	Introductory biology laboratory course	Each student collected data from PetFinder.com. Instructor summarised students' data for HWE analysis.	Y	Y	Y	This lab investigation highlights evidentiary reasoning that is essential for the entire investigation process. A practical guide provides scaffolding questions (Table 3) that can be implemented by instructors in their course.

Discussion

In this report, we applied the CADE framework in the context of our HWE laboratory class to illustrate how the implementation of the CADE framework influenced the laboratory investigation design, how we modified the lesson, as well as the actual changes in the laboratory instruction from one laboratory instructor as the lab discussions about the investigation were explicitly guided. The CADE framework targets known difficulties with students' evidentiary reasoning during investigations by giving them the opportunity to blend deep disciplinary knowledge and investigative practices with the epistemic considerations that are key to evidentiary reasoning. CADE also explicitly draws the instructors' attention to the distinction between disciplinary knowledge and epistemic considerations of evidentiary reasoning and how to prompt for each to promote students' deeper thinking about evidence during the investigation (Samarapungavan 2018). The scaffolding questions listed in Table 3 cover both aspects of disciplinary specificity and epistemic generality for three relationships to consider: *Theory => Evidence relationships* refer to the use of relevant disciplinary knowledge to inform what counts as evidence for a particular area of study, and what sort of data to collect; *Evidence <=> Data* relationships recognise the need for disciplinary knowledge to inform the isolation and definition of variables for a particular research design, procedures such as appropriate sampling and measurement, as well as the precision and accuracy of techniques and equipment to be used in order to optimise the reliability of the evidence; *Evidence => Theory relationships* involve the disciplinary knowledge implicit to the tracking and quantifying of known sources of error, alternate interpretations that could be evaluated, limitations and uncertainties to be explicitly addressed, and whether findings are consistent with disciplinary knowledge or raise questions about the ideas from previous studies that inform the investigation. Such details may not all need to be given in every laboratory protocol if the laboratory investigation provides opportunities to engage the students with discussions of these practices.

Guided by the CADE framework, an instructor was able to scaffold evidentiary reasoning in the process of the dog HWE investigation. This laboratory investigation uses a common pet as the subject to investigate, which was intended to motivate students for their laboratory investigation since many students love dogs. It calls on detailed disciplinary knowledge by providing students with information about the exact genes that control the phenotypes they can investigate, including *TYRP1* and *RSPO2* allele variants. After learning about *TYRP1* and *RSPO2* allele variants, most students can describe the relevant phenotypes and predict the genotype of a dog they know best, which allows students to focus on both the theory and scientific evidence underlying the HWE laboratory investigation. Students have an opportunity to connect disciplinary knowledge for their study to the detailed context of a familiar biological phenomenon that calls on their personal knowledge and experience, which is an essential component for development of evidentiary reasoning according to the CADE framework. The dog population investigation provides opportunities for students to practise hypothesis testing during this process to give purpose to the investigation as they gather and use evidence. The guided discussions about what variables are important for their research, how to collect data for use as evidence, what sampling strategies to use, and what conclusion(s) to draw from the evidence for the investigation, are intended to help students practise thinking about the nature, scope and quality of the evidence. From the teaching perspective, the embedded scaffolding questions presented in Table 3 are ready to apply for guiding students' evidentiary reasoning. Unlike other HWE activities that use coloured buttons, beads, or other manipulatives, the dog population investigation uses Petfinder.com, which is low cost and easy to access for implementation and thus provides a potential online laboratory teaching format for engaging students with authentic research.

HWE is a fundamental model that plays an essential role for understanding evolution in undergraduate biology education. Due to the abstract nature and mathematical foundation of HWE, students who mainly perform calculations with the HWE equation may not actually understand the biological phenomena underlying their investigations (Wise 2018b). Because of the importance as well as the challenges for teaching and learning HWE, educational researchers and instructors have been designing and implementing activities that aim to facilitate students' understanding as they apply the HWE in authentic and engaging ways (Bonner et al. 2019; Brewer and Gardner 2013; Smith 2017; Wise 2018a). However, there is room for improvements that could better engage students in thinking and reasoning with and about scientific evidence during the planning stages and not just for analysis of evidence from their HWE laboratory investigations. Research with the CADE framework could guide modification of other types of HWE laboratory investigations. For instance, the activities listed in Table 5 could be modified to give students more opportunities to practice reasoning about multiple aspects of scientific evidence. Modifications to a laboratory activity can be informed by the CADE framework to expand the activities from use of the Chi-square test to compare expected and observed allele frequencies in a population between different generations or in different places, by also giving students the opportunity to explore both disciplinary and epistemic components to evidence for changes in a population that could be attributed to natural selection or genetic drift. The Supplemental Material 1 provides a handout for teaching and learning HWE.

Others have reported that instructors' questions can engage and guide students through investigations by eliciting and scaffolding students' thinking if the instructors actually lead the discussions about observations, assumptions and reasoning (Kawalkar and Vijapurkar 2013). However, in an undergraduate laboratory where instruction is typically managed by graduate and undergraduate teaching staff who get very little pedagogical training, it can be challenging to get staff to help students discuss the evidence rather than just telling students what to do (Gardner and Gail Jones 2011; Luft et al. 2004; Sundberg, Armstrong, and William Wischusen 2005). Our findings show that after implementing the CADE framework to modify the lesson for laboratory instruction, a participant instructor used more questions related to multiple aspects of evidentiary reasoning for guiding her students through the HWE laboratory investigation. This suggests that the instructor would be able to lead students to think and reason more about the scientific evidence throughout the process of scientific investigation. For brevity, the detailed student data has not been presented here. Instead, we show how the CADE framework helped an instructor notice the importance of epistemic considerations in evidentiary reasoning. The instructor began to include scaffolding questions that prompted students to think and reason about the limitations and the nature, scope and quality of the evidence. Reasons for these interesting changes include the improvement in the laboratory task, but also, importantly, that the CADE framework helped unpack the complexity of evidentiary reasoning into questions that an instructor could implement to highlight essential components of evidentiary reasoning throughout the process of authentic research activity.

Limitations and directions for future research

There are several limitations to this study. First, we must avoid generalising our claims because the participant instructor may have been more excellent than a typical laboratory instructor in terms of both academic performance in research and teaching experience. Her laboratory instruction showed deep understanding of the CADE framework. There is an additional need to explore professional development approaches for laboratory instructors. With the CADE framework as an introduction to components of evidentiary reasoning and by providing both typical and refined lesson in terms of CADE-informed opportunities for students' evidentiary reasoning, it may be possible to help all laboratory instructors implement the CADE framework during all phases of a lab investigation to better guide discussions of evidence of relevance to their laboratory investigations. With the dynamic and interactive features of scaffolding, the goal of including evidentiary reasoning discussion could become more feasible. This study was focused on instruction of the HWE. Future research could examine how instruction informed by the CADE framework could influence and perhaps improve students' competence to use and reason with and about scientific evidence. Research is also needed to examine how and in what conditions the CADE framework should inform scaffolding questions to cultivate a learning environment and laboratory classroom atmosphere where students will become adept at thinking and reasoning with scientific evidence autonomously during their scientific investigations. By addressing these limitations, the development of evidentiary reasoning ability could benefit not only students' academic performance and science career preparation, but also strengthen their decision-making skills in everyday life, which is an important target for 21st century laboratory instructions.

Acknowledgements

We thank Dr. Trevor Anderson, Dr. Siddika SelcenGuzey and members of our research group focused on Exploring BiologicalEvidence for their contributions to the progress of our study. This study wassupported by the National Science Foundation under grant #1661124.Additionally, this work was informed by the ACE-Bio Network (NSF grant #1346567).Any opinions, findings, and conclusions or recommendations expressed in thismaterial are those of the authors and do not necessarily reflect the views ofthe National Science Foundation. We especially thank our laboratory class staffand students for willingly participating in this study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Science Foundation Grant/Award number 1661124.

ORCID

Chaonan Liu http://orcid.org/0000-0002-1786-5540
Dayna L. Dreger http://orcid.org/0000-0002-3218-6431
Shiyao Liu http://orcid.org/0000-0002-5092-782X
Ala Samarapungavan http://orcid.org/0000-0003-4016-9743
Stephanie M. Gardner http://orcid.org/0000-0003-3988-343X
Kari L. Clase http://orcid.org/0000-0003-2749-6440
Nancy Pelaez http://orcid.org/0000-0002-6328-5010

References

AAAS. 2011. "*Vision and Change in Undergraduate Biology Education: A Call to Action*." C. A. Brewer and D. Smith, Editors. Washington, D.C: American Association for the Advancement of Science (AAAS). https://visionandchange.org/finalreport/

Barzilai, S., and C. A. Chinn. 2018. "On the Goals of Epistemic Education: Promoting Apt Epistemic Performance." *Journal of the Learning Sciences* 27 (3): 353–389. doi:10.1080/10508406.2017.1392968.

Belland, B. R., K. D. Glazewski, and J. C. Richardson. 2008. "A Scaffolding Framework to Support the Construction of Evidence-Based Arguments among Middle School Students." *Educational Technology Research and Development* 56 (4): 401–422. doi:10.1007/s11423-007-9074-1.

Blumer, L. S., and C. W. Beck. 2019. "Laboratory Courses with Guided-Inquiry Modules Improve Scientific Reasoning and Experimental Design Skills for the Least-Prepared Undergraduate Students." *CBE—Life Sciences Education* 18 (1): 1. doi:10.1187/cbe.18-08-0152.

Bonner, K., D. Piechnik, J. Kovacs, A. Warwick, and P. White. 2019. "Clam Spawning & Red Tide: Helping Students Learn the Hardy-Weinberg Equilibrium." *The American Biology Teacher* 81 (5): 366–371. doi:10.1525/abt.2019.81.5.366.

Brewer, M. S., and G. E. Gardner. 2013. "Teaching Evolution through the Hardy-Weinberg Principle: A Real-Time, Active-Learning Exercise Using Classroom Response Devices." *American Biology Teacher* 75 (7): 476–479. doi:10.1525/abt.2013.75.7.6.

Cadieu, E., M. W. Neff, P. Quignon, K. Walsh, K. Chase, H. G. Parker, B. M. Vonholdt, et al. 2009. "Coat Variation in the Domestic Dog Is Governed by Variants in Three Genes." *Science* 326 (5949): 150–153. doi:10.1126/science.1177808.

Chinn, C. A., and B. A. Malhotra. 2002. "Epistemologically Authentic Inquiry in Schools: A Theoretical Framework for Evaluating Inquiry Tasks." *Science Education* 86 (2): 175–218. doi:10.1002/sce.10001.
Chinn, C. A., S. Barzilai, and R. Golan Duncan. 2020. "Education for a 'Post-Truth' World: New Directions for Research and Practice." *Educational Researcher* 50 (1): 51–60. doi:10.3102/0013189X20940683.
Christensen, A. C. 2000. "Cats as an Aid to Teaching Genetics." *Genetics* 155 (3): 999–1004. doi:10.1093/genetics/155.3.999.
David, W., J. S. Bruner, and G. Ross. 1976. "The Role of Tutoring in Problem Solving." *Journal of Child Psychology and Psychiatry* 17 (1): 89–100. doi:10.1080/0305498960220101.
Dobzhansky, T. 1973. "Nothing in Biology Makes Sense except in the Light of Evolution." *The American Biology Teacher* 35: 125–129. doi:10.2307/4444260.
Dreger, D. L., B. N. Hooser, A. M. Hughes, B. Ganesan, J. Donner, H. Anderson, L. Holtvoigt, K. J. Ekenstedt, and C. E. Ambrósio. 2019. "True Colors: Commercially-Acquired Morphological Genotypes Reveal Hidden Allele Variation among Dog Breeds, Informing Both Trait Ancestry and Breed Potential." *PLoS ONE* 10 14 doi:10.1371/journal.pone.0223995.
Duncan, R. G., C. A. Chinn, and S. Barzilai. 2018. "Grasp of Evidence: Problematizing and Expanding the Next Generation Science Standards' Conceptualization of Evidence." *Journal of Research in Science Teaching* 55 (7): 907–937. doi:10.1002/tea.21468.
Erduran, S., and M. Pilar Jiménez-Aleixandre. 2012. "Argumentation in Science Education Research: Perspectives from Europe." In *Science Education Research and Practice in Europe*, edited by D. Jorde and J. Dillon, 253–289. Rotterdam, The Netherlands: Sense Publishers.
Gardner, G. E., and M. Gail Jones. 2011. "Pedagogical Preparation of the Science Graduate Teaching Assistant: Challenges and Implications." *Science Educator* 20 (2): 31–41.
Gerber, B. L., A. M. L. Cavallo, and E. A. Marek. 2001. "Relationships among Informal Learning Environments, Teaching Procedures and Scientific Reasoning Ability." *International Journal of Science Education* 23 (5): 535–549. doi:10.1080/09500690116971.
Hannafin, M., S. Land, and K. Oliver. 1999. "Open Learning Environments: Foundations, Methods, and Models." In *Instructional-Design Theories and Models: A New Paradigm of Instructional Theory (Volume II)*, edited by C. Reigeluth, 115–140, New York: Routledge.
Hardy, G. H. 1908. "Mendelian Proportions in a Mixed Population." *Science* 28 (706): 49–50. doi:10.1126/science.28.706.49.
Hoffman, K., S. Leupen, K. Dowell, K. Kephart, J. Leips, and J. Batzli. 2016. "Development and Assessment of Modules to Integrate Quantitative Skills in Introductory Biology Courses." *CBE—Life Sciences Education* 15 (2): ar14. doi:10.1187/cbe.15-09-0186.
Jancuskova, T., M. Langevin, and S. Pekova. 2018. "TYRP1:C.555T>G Is a Recurrent Mutation Found in Australian Shepherd and Miniature American Shepherd Dogs." *Animal Genetics* 49 (5): 500–501. doi:10.1111/age.12709.
Jensen, J. L., and A. Lawson. 2011. "Effects of Collaborative Group Composition and Inquiry Instruction on Reasoning Gains and Achievement in Undergraduate Biology." *CBE—Life Sciences Education* 10 (1): 64–73. doi:10.1187/cbe.10-07-0089.
Jeong, H., N. B. Songer, and S. Young Lee. 2007. "Evidentiary Competence: Sixth Graders' Understanding for Gathering and Interpreting Evidence in Scientific Investigations." *Research in Science Education* 37 (1): 75–97. doi:10.1007/s11165-006-9014-9.
Kawalkar, A., and J. Vijapurkar. 2013. "Scaffolding Science Talk: The Role of Teachers' Questions in the Inquiry Classroom." *International Journal of Science Education* 35 (12): 2004–2027. doi:10.1080/09500693.2011.604684.
Kinnear, J. F. 1986. "Using the Domestic Cat in the Teaching of Genetics." *Journal of Biological Education* 20 (1): 5–11. doi:10.1080/00219266.1986.9654766.
Kuhn, D. 2009. "Do Students Need to Be Taught How to Reason?" *Educational Research Review* 4 (1): 1–6. doi:10.1016/j.edurev.2008.11.001.
Labrinidis, A., and H. V. Jagadish. 2012. "Challenges and Opportunities with Big Data." *Proceedings of the VLDB Endowment* 5 (12): 2032–2033. doi:10.14778/2367502.2367572.
Laursen, S. 2019. *Levers for Change: An Assessment of Progress on Changing STEM Instruction*. Washington, DC: American Association for the Advancement of Science. https://www.aaas.org/resources/levers-change-assessment-progress-changing-stem-instruction
Luft, J. A., J. P. Kurdziel, G. H. Roehrig, and J. Turner. 2004. "Growing a Garden without Water: Graduate Teaching Assistants in Introductory Science Laboratories at a Doctoral/Research University." *Journal of Research in Science Teaching* 41 (3): 211–233. doi:10.1002/tea.20004.
Manz, E. 2016. "Examining Evidence Construction as the Transformation of the Material World into Community Knowledge." *Journal of Research in Science Teaching* 53 (7): 1113–1140.
Marx, V. 2013. "The Big Challenges of Big Data." *Nature* 498 (7453): 255–260. doi:10.1038/498255a.
Masel, J. 2012. "Rethinking Hardy-Weinberg and Genetic Drift in Undergraduate Biology." *BioEssays* 34 (8): 701–710. doi:10.1002/bies.201100178.

Next Generation Science Standards Lead States. 2013. *Next Generation Science Standards: For States, by States*. Washington DC: National Academies Press. https://ole.sandiego.edu/bbcswebdav/pid-1217592-dt-content-rid-10695894_1/courses/EDUC-509-MED-MASTER/M4/M4_NGSS_App_I.pdf

Osborne, J., S. Erduran, and S. Simon. 2004. "Enhancing the Quality of Argumentation in School Science." *Journal of Research in Science Teaching* 41 (10): 994–1020.

Pelaez, N., T. Anderson, S. Gardner, Y. Yin, J. Abraham, E. Bartlett, C. Gormally, et al. 2017. "The Basic Competencies of Biological Experimentation: Concept-Skill Statements." *PIBERG Instructional Innovation Materials*. *Paper 4*, no. 2017. http://docs.lib.purdue.edu/pibergiim/4

Pelaez, N., S. M. Gardner, and T. R. Anderson. 2022."The Problem with Teaching Experimentation: Development and Use of a Framework to Define Fundamental Competencies for Biological Experimentation Pelaez, N., Gardner, S. M., Anderson, T. R.." In Trends in Teaching Experimentation in the Life Sciences: Putting Research into Practice to Drive Institutional Change. Switzerland: Springer Nature 3-27.

Puntambekar, S., A. Stylianou, and J. Goldstein. 2007. "Comparing Classroom Enactments of an Inquiry Curriculum: Lessons Learned from Two Teachers." *The Journal of the Learning Sciences* 16 (1): 81–130.

Samarapungavan, A. 2018. "Construing Scientific Evidence: The Role of Disciplinary Knowledge in Reasoning with and about Evidence in Scientific Practice." In *Scientific Reasoning and Argumentation*, edited by F. Fischer, C. Chinn, K. Engelmann, and J. Osborne, 56–76. New York: Routledge.

Sandoval, W. A., and B. J. Reiser. 2004. "Explanation-Driven Inquiry: Integrating Conceptual and Epistemic Scaffolds for Scientific Inquiry." *Science Education* 88 (3): 345–372. doi:10.1002/sce.10130.

Sandoval, W. A., and K. A. Millwood. 2005. "The Quality of Students' Use of Evidence in Written Scientific Explanations." *Cognition and Instruction* 23 (1): 23–55. doi:10.1207/s1532690xci2301_2.

Sarit, B., and C. A. Chinn. 2020. "A Review of Educational Responses to the 'Post-Truth' Condition: Four Lenses on 'Post-Truth' Problems." *Educational Psychologist* 55 (3): 107–119. doi:10.1080/00461520.2020.1786388.

Schmutz, S. M., T. G. Berryere, and A. D. Goldfinch. 2002. "TYRP1 and MC1R Genotypes and Their Effects on Coat Color in Dogs." *Mammalian Genome* 13 (7): 380–387. doi:10.1007/s00335-001-2147-2.

Schmutz, S. M., and T. G. Berryere. 2007. "Genes Affecting Coat Colour and Pattern in Domestic Dogs: A Review." *Animal Genetics* 38 (6): 539–549. doi:10.1111/j.1365-2052.2007.01664.x.

Schuchardt, A. M., and C. D. Schunn. 2016. "Modeling Scientific Processes with Mathematics Equations Enhances Student Qualitative Conceptual Understanding and Quantitative Problem Solving." *Science Education* 100 (2): 290–320. doi:10.1002/sce.21198.

Shim, S. W., and N. Pelaez . 2022. "Getting by with A Little Help from Friends: A Qualitative Case Study of Students'Strategies for Coping with Failure in an Undergraduate Biology Laboratory Course." *CBE—Life Sciences Education* 21 (2): ar17, 1–21. doi:10.1187/cbe.20-07-0155.

Smith, M. U., and J. T. Baldwin. 2015. "Making Sense of Hardy-Weinberg Equilibrium." *American Biology Teacher* 77 (8): 577–582. doi:10.1525/abt.2015.77.8.3.

Smith, M. U. 2017. "Teaching the Hardy-Weinberg Equilibrium: A 5E Lesson Plan." *American Biology Teacher* 79 (4): 288–293. doi:10.1525/abt.2017.79.4.288.

Speth, E. B., J. L. Momsen, G. A. Moyerbrailean, D. Ebert-May, T. M. Long, S. Wyse, D. Linton, and J. Jungck. 2010. "1, 2, 3, 4: Infusing Quantitative Literacy into Introductory Biology." *CBE—Life Sciences Education* 9 (3): 323–332. doi:10.1187/cbe.10-03-0033.

Stern, C. 1943. "The Hardy-Weinberg Law." *Science* 97 (2510): 137–138. doi:10.1126/science.97.2510.137.

Sundberg, M. D., J. E. Armstrong, and E. William Wischusen. 2005. "A Reappraisal of the Status of Introductory Biology Laboratory Education in U.S. Colleges & Universities The American Biology Teacher 67 (9): 525–529 ." doi:10.2307/4451904.

Tabak, I., and E. Baumgartner. 2014. "The Teacher as Partner: Exploring Participant Structures, Symmetry." *And Identity Work in Scaffolding." Cognition and Instruction* 22 (4): 393–429.

Ushiri, K., R. S. Moog, and J. E. Lewis. 2013. "Argumentation and Participation Patterns in General Chemistry Peer-Led Sessions." *Journal of Research in Science Teaching* 50 (10): 1207–1231. doi:10.1002/tea.21107.

van Joolingen, W. R., and T. De Jong. 1991. "Supporting Hypothesis Generation by Learners Exploring an Interactive Computer Simulation." *Instructional Science* 20 (5–6): 389–404. doi:10.1007/BF00116355.

Vygotsky, L. S. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, Mass: Harvard University Press.

Weinberg, W. 1908. "Über Den Nachweis Der Vererbung Beim Menschen." *Jahres. Wiertt. Ver. Vaterl. Natkd* 64: 369–382. doi:10.1007/BF00531038.

Wilson, C. D., J. A. Taylor, S. M. Kowalski, and J. Carlson. 2010. "The Relative Effects and Equity of Inquiry-Based and Commonplace Science Teaching on Students' Knowledge, Reasoning, and Argumentation." *Journal of Research in Science Teaching* 47 (3): 276–301. doi:10.1002/tea.20329.

Wise, M. J. 2018a. “How to Teach the Hardy-Weinberg Principle Using Engaging, Non-Trivial Evolutionary Scenarios.” *Bioscene* 44 (1): 28–36.
Wise, M. J. 2018b. “On Hardy-Weinberg in Introductory Biology: Teaching This Fundamental Principle in an Authentic, Engaging, and Accurate Manner.” *Bioscene* 44 (1): 37–41.
Xun, G., and S. M. Land. 2003. “Scaffolding Students’ Problem-Solving Processes in an Ill-Structured Task Using Question Prompts and Peer Interactions.” *Educational Technology Research and Development* 51 (1): 21–38. doi:10.1007/BF02504515.
Zarocostas, J. 2020. “How to Fight an Infodemic.” *The Lancet* 395 (10225): 676. doi:10.1016/S0140-6736(20)30461-X.