Under the Microscope:
Review of the research on biological lab experiences 1987-2007
A Research White Paper

by

Brian Drayton[1], Gillian Puttick, Meaghan Donovan

TERC

[1] Contact: brian_drayton@terc.edu, gilly_puttick@terc.edu

1. INTRODUCTION

A sound education in the fundamentals of biology is increasingly important for all citizens of the 21st century (Kress and Barrett 2001). Issues ranging from environmental change to medical genetics require a populace that understands not just "biology facts," but also the nature of biological investigation and evidence, and the evolutionary basis that underlies the entire field. Many crucial developments in our world cannot be well understood without some facility with the kind of reasoning about complex, variable, and hierarchical systems that are typical of biology. Biology education also plays an important "gatekeeper" role for the sciences. After "general science" (which often includes a life science component), biology is the discipline that the highest proportion of 17-year-olds have taken, and thus may be the final science many students encounter (NAEP 1999).

Practical engagement with problems that possess these living characteristics is an indispensable ingredient in the development of this kind of reasoning and understanding. Laboratory experiences[1] have been seen as essential parts of high-quality biology education since the 19th century (Lurie, 1988; Deboer, 1991; Janovy, 2003), as a way to convey biology content, and to help students understand how biology is done. However, the actual contribution of laboratories to student learning remains actively open to question (NRC, 2006).

A recent groundbreaking study, *America's Lab Report* (National Research Council [NRC] 2006), has identified significant "frontiers of ignorance" awaiting attention; their findings echo those of an earlier survey of the role of laboratory experiences in science by Lazarowitz and Tamir (1994). For example, at the outset of the review, *America's Lab Report* listed several learning goals which high school laboratory experiences are assumed to further. However, the NRC survey of the US research literature left the authors asking such basic questions as "What are the specific learning outcomes of laboratory experiences? What are the teaching and learning processes by which laboratory experiences contribute to particular learning outcomes for diverse learners and different populations of students? What kinds of curriculum can support teachers and students in progress toward these learning outcomes?" (p. 10). Furthermore, the NRC study limited its research to high school, expressly leaving aside the question of earlier grade levels, and the progression from elementary school to college.

Finally, and most important for our purposes, the NRC study does not explore any of these questions from the point of view of particular disciplinary content. While some studies of high school biology lab experiences are referenced in their work, the authors did not, even within the chosen limitations of their study, set out to systematically review the literature on the life sciences. Thus, some of the unique characteristics of the biological sciences have not been used in a consideration of the goals, value, and growth of understanding in biology during students' biology education careers.

This paper reports results from a study, funded by the National Science Foundation[2], which examined the nature of published research on life-science laboratory experiences in the English-language, peer-reviewed research literature. The study sought to understand the extent to which the research literature addressed i) student learning about the characteristics of living systems, ii) the growth of students' biological reasoning with respect to "biological distinctives" (evolution, complexity, dual causation, populational thinking, and variability) (Mayr 1982, 2004; Sterelnyi and Griffiths, 1999), and iii) novel or standard lab activities.

We examined research methodology, biological subject matter addressed, and materials and methods used in the lab. We sought to establish some estimate of the degree of student inquiry that the lab allowed. This paper will focus on our findings on: i) the size and basic characteristics of the corpus, ii) the biological domain addressed, iii) the degree of inquiry, pedagogical design and activity structure of the lab experiences studied, and, iv) student outcomes measured in the studies reported.

Summary findings

We can very briefly summarize our findings as follows: Based on our initial investigation of the literature, as described above, we made a series of hypotheses about what we would find. We conjectured that the research base:

1. Includes more studies in some grade levels than others. The evidence of our study strongly supports this. See §3D.
2. Fails to cover many biology topics. This hypothesis is also strongly supported by this study. See §6.
3. Emphasizes innovative practices or experiments rather than focusing on common practice. This innovation bias is very strong, on the evidence from this corpus, see §4E.
4. Does not show much change in basic activity structure of lab experiences from elementary grades to college. The evidence on this hypothesis is inconclusive. We do not believe one could argue against it on the basis of the studies we examined, but very few of the studies provide enough detail about the labs to allow for a judgment. See §7.
5. Does not provide evidence about how laboratory experiences relate to students' understanding of the distinctive features of biological systems. The study does provide some evidence on which to judge this; a full analysis is found in §9.
6. Shows little variation in the complexity of systems used for the investigations. This hypothesis is supported. see §7A.
7. Does not allow us to address the question of progressions, in particular bearing in mind the distinctive characteristics of biology. This hypothesis is supported, see §8
8. Provides fragmentary evidence about student outcomes, with a wide range of outcome measures used, variation in inferential rigor provided by the study methodology, and characterizations of interventions being studied that are uneven in descriptive adequacy. See §10.

Background

The motivation for our study lies in the challenges now facing life science education. Evidence from recent studies about students' declining interest in science is alarming, especially in light of a rising need for public understanding of biological science in particular, and a concomitant demand for more researchers in the life sciences (including in critical understaffed fields such as systematics) (NCES 2006, National

Science Board 2006). Such considerations seem to demand a re-evaluation of biology teaching and learning. An examination of the role of "practical work," or laboratory experiences as characterized by the National Academies, must form one necessary element of this reevaluation. A first step in such analysis is to ask, "What does the literature say?" Without a synthesis of findings in the published literature, we can not answer that question.

The need for this study can be simply put, therefore: No systematic review of the recent literature on life science lab experiences exists. This bald statement is based on a broad search in relevant publications. Indexes and tables of contents for prominent journals were examined for the period under review. These journals include *Science Education*, Bioscience, *Journal of Research in Science Teaching*, *International Journal of Science Education*, *Review of Educational Research* (AERA), *Review of Educational Research* (Australia), and *Canadian Journal of Educational Research*. When articles on biology labs were found, their references were checked for possible review articles referred to. This journal-by-journal search was complemented by web-searches (using keywords such as *biology* or *life science* and *laboratory*, *synthesis*, *analysis,* and *meta-analysis)*, and by a search of the ERIC and JSTOR databases. Finally, three major recent books which would be likely to make use of such a review were checked for references: The 2007 *Handbook of Science Education* (Abell and Lederman 2007), and the two National Research Council studies, *Taking science to school (NRC 2007)* and *America's lab report NRC 2006)*. This search resulted in the discovery of articles reporting on experimental or observational studies, as well as numerous derivative articles, but no syntheses of research.

This report is designed to render several benefits for the field as a whole, and for our own future research and development. It is intended to:

- Identify scholarly consensus where it exists on the values, problems, and implications of particular kinds of laboratory experiences
- Enable the field to fill identified gaps in research about the role of laboratory experiences in life-science learning
- Contribute a theoretical framework for the study of lab experiences in life-science education, and provide a foundation for studies of learning and teaching issues arising from distinctive characteristics of living systems and their study
- The framework thus derived is also a valuable pre-requisite for the future analysis of curriculum materials for life-science laboratory experiences with respect to some of the key parameters outlined for the proposed study.

As argued in *America's Lab Report* (NRC 2006), there is an urgent need for practical, challenging, rigorous, and inquiry-rich encounters in labs across the spectrum from elementary grades to college to ensure that students encounter actual science practice in the classroom, and are helped systematically to acquire a basic grasp of biological practice and understanding. Science standards and other frameworks, such as NRC (1996) and AAAS (1993) and the many state standards developed in the past decade, reiterate this need, but provide limited guidance about when to incorporate what kinds of laboratory experiences and for what purpose.

Furthermore, this kind of guidance is most helpful when it is *specific* to the particular phenomena and concepts being worked with and the particular investigative methods and skills associated with these (NRC 2007; Janovy 2003). The time is ripe for focused research on biology pedagogy, and this significantly includes the role and effectiveness of laboratory experiences. Such research is essential to serve as a basis for a future program of research-based curriculum change and teacher professional development.

In order that such research should build on firm results already achieved, and address unexplored terrain, the current landscape of research on biology lab experiences needs to be examined critically, within a coherent framework that includes understandings about classroom practice, cognitive development, the role of state standards, core competencies valued by life scientists, and the nature of biology. A survey that makes available "lessons learned to date" is the first step for such a program, which will involve critical examination of current materials and practice. We suggest that such an examination should, while bearing in mind general values and approaches shared by the sciences, take explicit account of features that are unique or highly characteristic of biological systems, and therefore of the reasoning and methodologies that are appropriate to these distinctive features. It is here that the philosophy of biology provides useful insights, which can be used to interrogate both the practice of teaching biology and life sciences, and the literature that studies such practices — most especially the "practical work" of laboratory experiences and hands on activities.

2. METHODOLOGY

A. Article acquisition and size of corpus

We sought English-language research on biology lab experiences that have a clearly articulated theoretical framework and research methodology, from the period 1988-2007, the period dominated by the development of national and state science and inquiry standards. We began seeking articles by conducting keyword searches of electronic versions, as being the most likely approach to provide a quick compilation of items for review. However, after doing such searches on 3 journals, we tested the results by inspecting the tables of contents and article abstracts for the journals searched. Because we found that even a very full set of keywords did not catch potential items, we decided to abandon the use of electronic searches, and instead read journal tables of contents and abstracts systematically for candidate articles. We established a list of 25* journals (included in Appendix).

Articles that reported research on life sciences laboratory experiences, from the appropriate grade levels (elementary through first year college) were acquired in hard copy (and if possible also in soft copy). All articles, when acquired, were given an accession number, and the bibliographical information was entered into an EndNote database. Soft copy, if available, was also stored in that database. Two copies were printed and labeled with the acquisition number, and filed.

Dissertations were sought from ProQuest Dissertations and Theses (Full Text) dissertation repository. Candidates were selected first on the basis of their abstracts. Copies of candidate dissertations were obtained either in hard or soft copy. Selected dissertations were referenced in the EndNote database.

Each article or dissertation was assigned to two researchers for coding. Coding was done on hard copy, with marginal notes identifying data used to assign codes. The coding for each item was recorded on coding sheets developed for the study. A record of article assignments and coding status was maintained in an Excel database.

B. The corpus used for this study

We found 211 articles and 20 dissertations that met our initial criteria as candidate items for coding. We examined the corpus of 231 candidate items for further inclusion/exclusion based on two principal criteria:

1) Was biology learning or teaching the focus of the article? Did the study context include a theoretical framework or explicit learning theory related to biology? While this was usually straightforward to determine, in a surprising number of cases it appeared that the researcher, studying phenomena unrelated to biology, such as the use of technology or efficacy of group learning, chose a life-science classroom as the context for the research. In response to problematic cases, we developed a four-part heuristic filter. When the focus/context was ambiguous, we answered the following questions:

1. Is biology explicitly included in the title?
2. Is biology explicitly included in the research question(s)?
3. Does the introduction or theoretical framework provide evidence that biology subject matter is integral to the study?
4. Are at least some of the reported outcomes explicitly relevant to biology?

If 3 of these could be answered affirmatively, the study continued to be considered for inclusion. Of the original 231 candidate items, we excluded 99 on the basis of this criterion.

2) Was the study's methodology described in sufficient detail to substantiate conclusions? We wanted to be able to identify how the conclusions were reached, what the study population was, data collection and analysis methods, and similar features of study design. Of the original 231 candidates, 22 were excluded on this criterion.

Our final corpus consisted of 110 studies.

C. Coding

Two researchers coded all articles; inter-rater reliability was above 75%. Where disagreements occurred, the coders discussed the differences, and established an agreed coding. Codes addressed key elements of study design and methodology, content addressed, nature of learning activity, materials used, presence or absence of "biological distinctives," degree of "directedness" of the lab, and student and other outcomes reported. A sample coding sheet is provided in the Appendix.

Once each article was coded, the codes were entered into a FilemakerPro database to facilitate analysis. For the present white paper, all included articles were examined and analyzed with respect to the research questions. An initial draft was completed by one researcher, and the analysis was then critiqued by two others; drafts of the analysis were presented to the project's advisory board for further critique.

3. BASIC CHARACTERISTICS OF CORPUS

A. Nationality

We were interested in whether or not the English-language research literature from different countries and/or regions might emphasize key conceptual or pedagogical issues that may be of value well beyond the country of origin. Interestingly enough, about half the papers in our corpus were from the United States. The remaining papers were scattered about other geographic locations, with only a handful of papers from each location (Fig. 1). This fragmentation of the corpus makes it difficult to identify trends by geographic location since the sample size for all regions, other than the US, are quite small. We may also have missed important research from Europe, Asia, the Middle East and Latin America published in non-English publications.

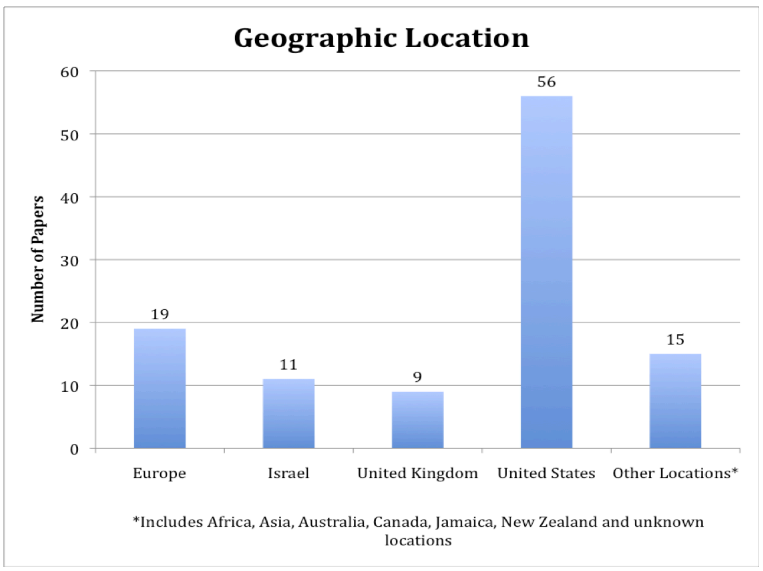


Figure 1.

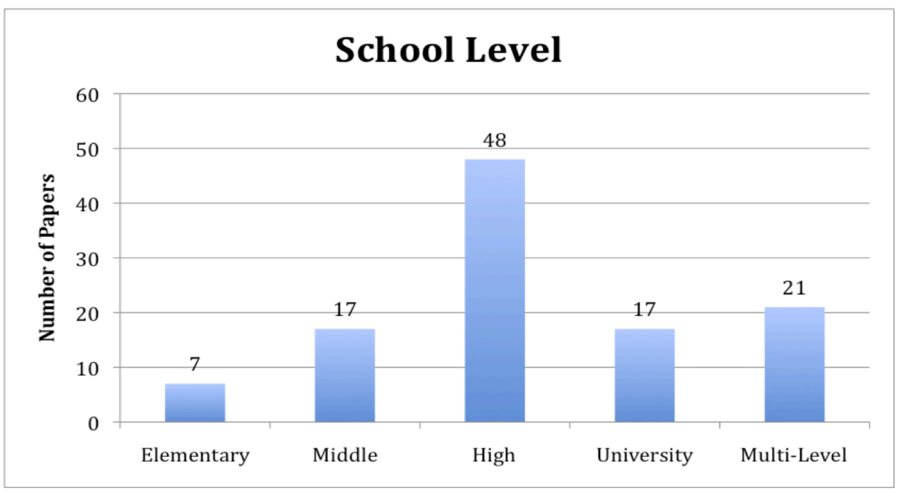
For the most part there where no major differences among the studies from different geographic regions. The few differences that did arise are described in detail below. The small sample size of most geographic regions also made it difficult to discern trends in the data.

B. Grade Level.

The most frequently reported educational level was high school (44%; 48 studies); middle school and university each constituted 17%, and elementary school 6%. (Note that for college, our synthesis only included entry-level biology courses, thus anything above grade 13 was not included). Nineteen percent included a mix of school levels. Twenty-one papers (19%) had study populations that included a mix of school levels; these are reported as Multi-Level.

International papers that reported on schools in which grade levels are labeled according to a system different from that in the US (e.g. 4th form in Nigeria) were translated into a US equivalent. Some papers were included in the Multi-Level category because of the way the school level categories were assigned. Out of the 21 papers in the Multi-Level category, 7 papers (33%) were placed there because the two-year range of the ages or grades reported in the paper fell between the cut-off ages for the different grade levels. For example, one paper reported studying students in the fifth and sixth grades. Since, for the purposes of this study, we defined elementary school as K-5 and middle school as 6-8, a study of fifth and sixth graders falls into both elementary and middle school categories. Similarly, another paper reported that the students in their study population were between the ages of 13 and 14. Once again, the definition set for the school levels in terms of age middle school (11-13) and high school (14-17), causes the study to span two school levels.

The remaining 14 papers (67%) out of 21 categorized as Multi-Level, included age or grade ranges that spanned three or more years/grades and fell within at least two different school level categories.



Elementary = grades k-5, Middle = grades 6-8, High = grades 9-12, University = grade 13

C. Ethnicity and Gender

Only papers from the United States reported on the ethnicity of their study groups with any consistency (34% of papers from the US reported on ethnicity). Israel, UK and Other Geographic Regions did not report any information on the ethnicity of their study subjects. Five percent of the papers from Europe reported on ethnicity.

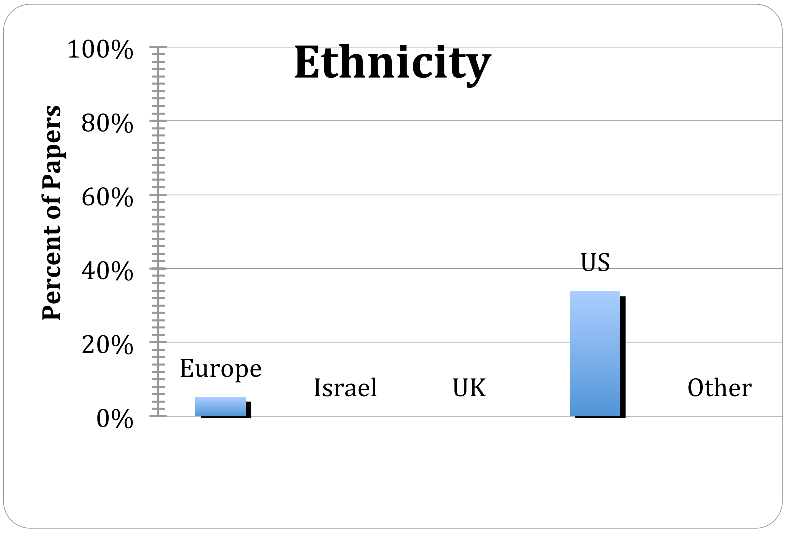


Figure 2.

It is also interesting to note that the reporting on the gender of study subjects was considerably lower in the United Kingdom (11% of papers), as compared to the other geographic locations with in which 36 – 53% of the papers in a given geographic location included gender data.

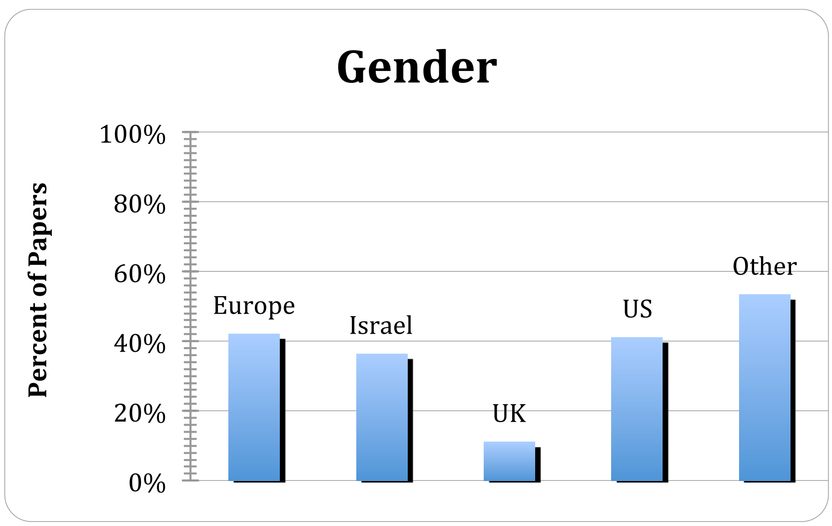


Figure 3.

D. Year of publication

We limited our search for research articles between the years of 1988 and 2007. Almost half of the papers (50 papers) were published within the last five years, possibly reflecting an increase in attention to lab activities during this time. An additional 34% (37 papers) were published between the dates of 1998 and 2002. The remaining 23 were published before 1997. (Figure 4)

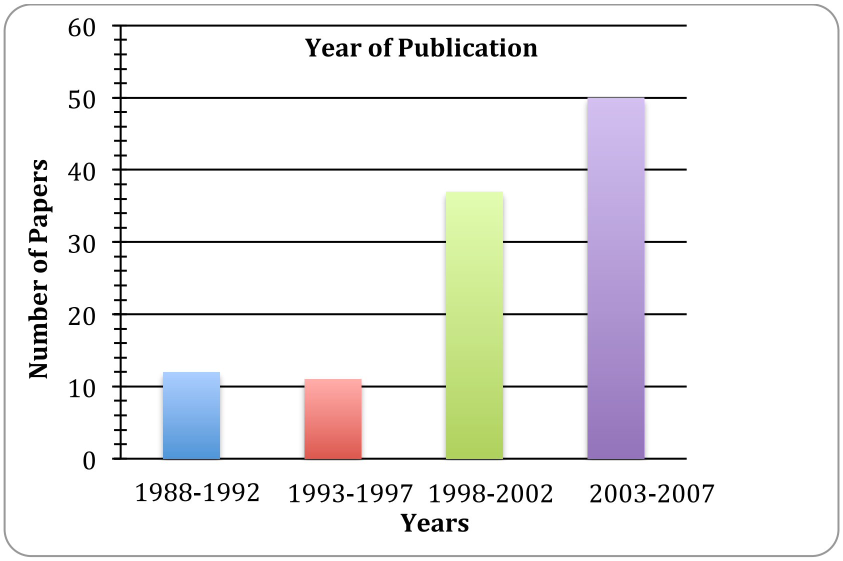


Figure 4.

4. DESIGN OF THE STUDIES

A. Study methodologies

Almost half (45%) of the studies used an experimental design that included control and intervention groups. In addition, there were 7 observational studies and 7 case studies, while the remaining 40 papers employed mixed methods designs. Six studies used more than one type of research methodology (Figure 5).

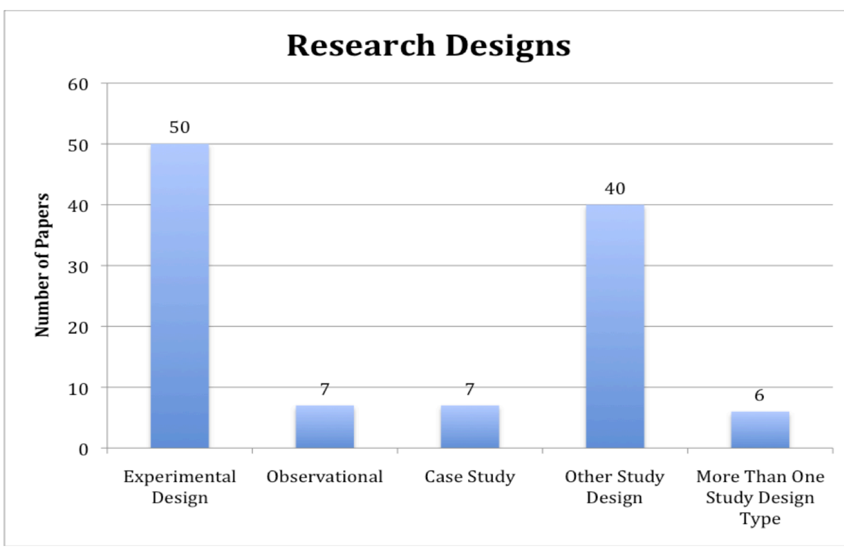


Figure 5.

B. Data Collection Methods

22 papers (20%) reported using only qualitative data collection methods, while 38 papers (35%) used only quantitative data collection methods. Fifty papers (45%) reported using a mix of both qualitative and quantitative data collection methods, or reported 'mixed-methods' as their type of data collection.

C. Number of students, instructional setting, and grade level

The number of students in the studies ranged from 6 to 4,000 (Figure 6). Fifty percent had study populations of 100 students or fewer while 46% included more than 100 students. Four percent did not report the size of the study population.

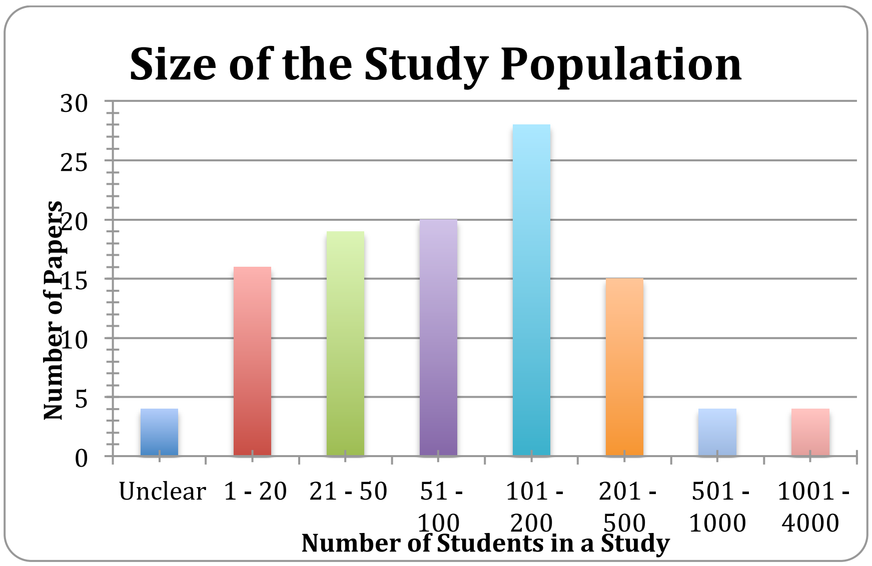


Figure 6.

Fifteen percent focused on a single class of students, 28% on multiple classes within a single school, and 32% were conducted in classes from multiple schools (Fig. 7). Eighteen percent brought a group of students together for the purpose of conducting their research.

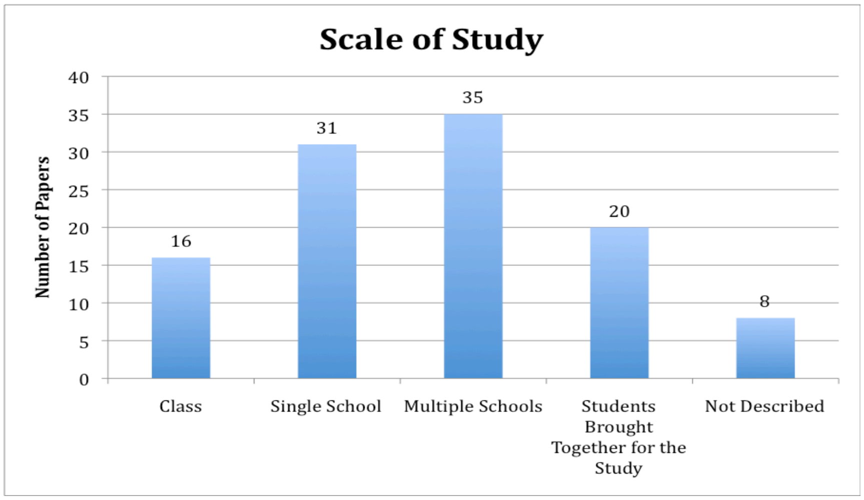


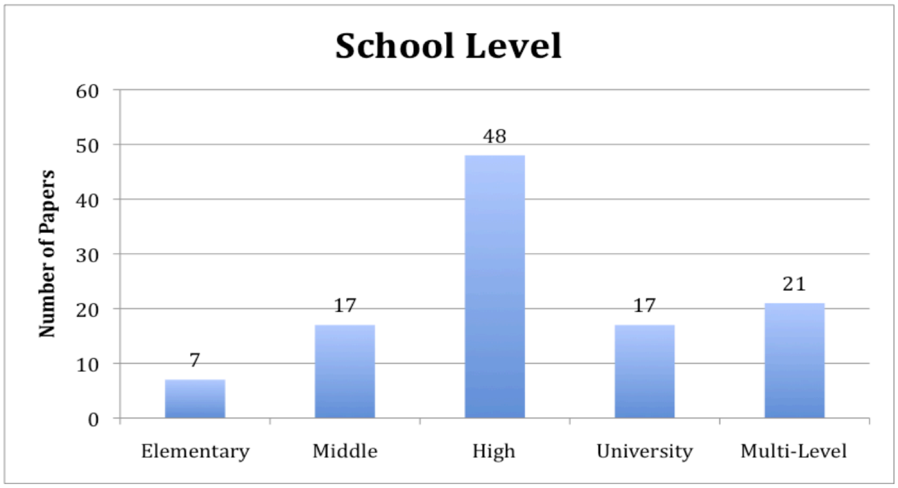
Figure 7.

D. Educational (Grade) level

The most frequently reported educational level was high school (44%; 48 studies); middle school and university each constituted 17%, and elementary school 6% (Fig. 8). (Note that for college, our synthesis only included entry-level biology courses, thus anything above grade 13 was not included). Nineteen percent included a mix of school levels. Twenty-one papers (19%) had study populations that included a mix of school levels; these are reported as Multi-Level.

International papers that reported on schools in which grade levels are labeled according to a system different from that in the US (e.g. 4th form in Nigeria) were translated into a US equivalent. Some papers were included in the Multi-Level category because of the way the school level categories were assigned. Out of the 21 papers in the Multi-Level category, 7 papers (33%) were placed there because the two-year range of the ages or grades reported in the paper fell between the cut-off ages for the different grade levels. For example, one paper reported studying students in the fifth and sixth grades. Since, for the purposes of this study, we defined elementary school as K-5 and middle school as 6-8, a study of fifth and sixth graders falls into both elementary and middle school categories. Similarly, another paper reported that the students in their study population were between the ages of 13 and 14. Once again, the definition set for the school levels in terms of age middle school (11-13) and high school (14-17), causes the study to span two school levels.

The remaining 14 papers (67%) out of 21 categorized as Multi-Level, included age or grade ranges that spanned three or more years/grades and fell within at least two different school level categories.



Elementary =grades k-5, Middle =grades 6-8, High=grades 9-12, University = grade 13

Figure 8.

Thirty-seven papers reported 'school classroom' as the setting in which the students they were studying conducted their laboratory activity (Fig. 9). Thirteen papers reported the setting as 'school lab,' eleven papers reported 'computer lab.' Nine papers reported that students conducted laboratory activities at a university or research institute. Fifteen papers indicated that the students they were studying went on a field trip. Thirteen papers reported that the students they were studying conducted laboratory activities during 'out of school time.' Eight papers reported a different setting than the choices described above. Thirty-one papers did not include information about where the students in the study conducted the laboratory activity. (Note that since 24 out of the 110 papers reported more than one setting, numbers in the chart total more than 110.)

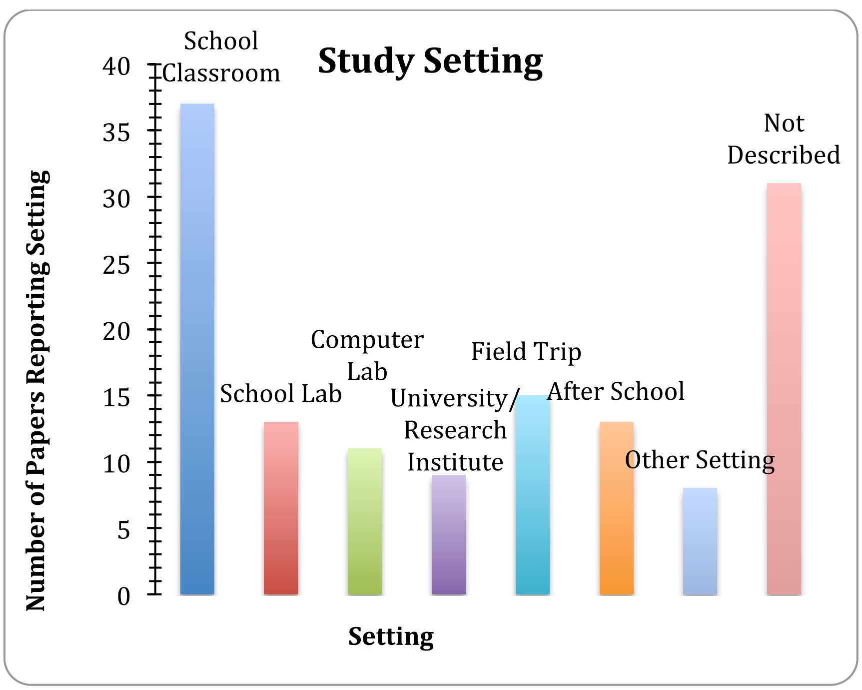


Figure 9.

E. Innovation bias

The papers in our study overwhelmingly report on innovative methods or materials (Figure 10). Over three quarters of the 110 coded papers, (86 papers) focused their research on a novel approach to a laboratory activity, while 17 (15%) focused on a standard or common laboratory activity. Only 3% researched both novel and standard labs to compare their efficacy.

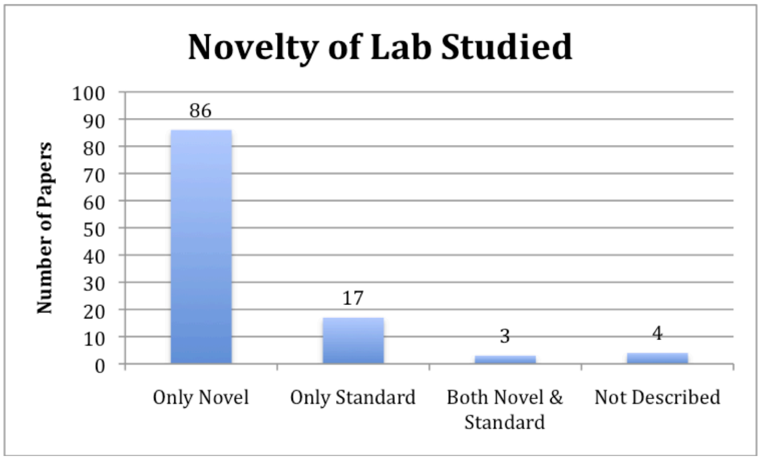


Figure 10.

Thus, the corpus of research literature in our study does not provide much information on the effectiveness of the kinds of activities that most students most often encounter. Generally, the papers focused on innovations in technology or technique do not build their rationales on the basis of evidence that a common, mainstream, or "traditional" lab activity has been shown to be ineffective in supporting learning. Instead, two common strategies are to use a standard lab as the comparison group in a study of an innovation, or to propose an activity (often involving a computerized system of some kind) as an alternative to traditional instruction where labs are not typically done.

For example, Hickey et al. (2003) describe a genetics simulation environment (GenScope) being used to address issues in genetics learning and reasoning in high school biology. The general form of the argument is, Can this environment help students learn genetic reasoning better than students instructed in other ways? The "other ways" are not analyzed, but used as a comparison group. The study shows real gains in some populations, and gives evidence that the GenScope environment thus has real promise as an augmentation of typical instruction.

In similar fashion, Geraedts and Boersma (2006) argue effectively that, whatever the methods used, teaching about evolution and natural selection in schools generally yields unsatisfactory results. Their paper describes a two-lesson "guided reinvention" of Darwinian theory including a critique of Lamarkian evolution. It is probably true that most secondary biology courses do not devote this much concentrated and designed attention to the development of the reasoning behind evolutionary theory. The authors make a strong case that doing so can improve students' understanding of concepts that need to be better taught; they are not suggesting that this activity is a replacement for a typical lab experience that has been shown to be ineffective.

Huppert et al. (1998) report on an experimental study that compared the learning about the growth curve of bacterial colonies by students who used a computerized simulation to augment other methods of instruction with students who had "ordinary" lab experiences. The students who used the computer environment showed significant gains in understanding, as compared with the students not using these tools. The paper shows how difficult it can be to compare apples and oranges, that is, very different kinds of experiences. While the study's methodology provided a fairly rigorous parallel structure for the experimental and the comparison groups, the paper states that "the students in this [control] group spent the same amount of time studying the learning unit." But the students using the simulation software were able to run replicate experiments, and vary the design.

While the comparison group conducted laboratory experiments, the characteristics of actual bacterial cultures meant that preparing and running multiple trials and varying designs to an extent comparable to the simulations was not possible within the time available. The simulations thus produced a quantity and quality of (simulated) evidence that enhanced students' grasp on the concepts and phenomena being addressed. Yet the control students were not given the opportunity to perform a similarly rich array of experiments, and did not have a comparably rich experience. One cannot say, therefore, how students' learning in similar circumstances would compare with that of the control group nor the experimental group in this present study.

5. DIVERSITY OF LEARNERS

A. Overview

Only 51% of the studies described student demographics (Fig. 11). We found this result surprising, given that educational research has raised many questions about the impact of cultural context on science learning (Lee and Louckx, 2000; Warren et al., 2001). Forty-four papers (40%) included information on the ratio of boys to girls, 21 (19%) on ethnicity of their study population, and 5% special populations, visually impaired students, or gifted students. More than half (53%) did not describe the socioeconomic status (SES) of their study population nor whether the study population was urban, suburban or rural.

Nearly half (49%) the studies included in the 110-paper corpus did not report information on the gender, ethnicity or special population (i.e. special education classes, visually impaired, etc.) of their study population. Most of the papers (43 papers / 39% of coded papers) that did report on gender, ethnicity and special population only included one category of demographics (e.g. the paper included ethnic data, but did not include information on gender or special population). Only 13 papers (12%) reported more than one category of demographics (e.g. the paper reported both the gender and the ethnic data of their study population).

The majority of the papers that included information on the demographics of their study population did not make further use of this information: gender, SES, and ethnicity were rarely included as a variable in data analysis. As will emerge from the data described in this section, the corpus provided remarkably little information about how the lab experiences being studied affected various groups of interest.

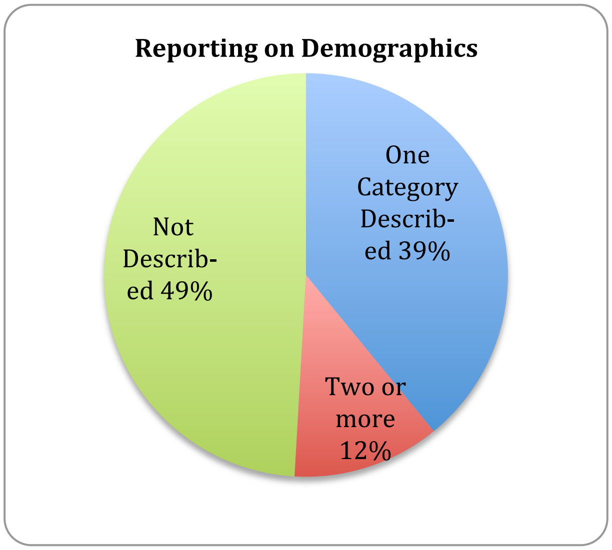


Figure 11.

B. Gender

Gender was the most common demographic variable reported by the studies in the corpus, with 44 of the 110 papers (40%) including information on the ratio of boys to girls. The remaining 66 papers (60%) did not report information on the gender of the study population.

Twenty-five papers merely quoted the proportions of males and females in their study population but did not use these data in their analysis of results. Seven of the 44 papers that reported information on gender used the information to balance gender representation in their study population in some way. Five of the seven studies selected even numbers of boys and girls for their studies. For example, in their 2004 paper of conceptual change in learning genetics Tsui and Treagust selected two boys and two girls to participate in their study. The remaining two studies did not select even numbers of boys and girls to participate in the study, but balanced gender in other ways. Soyibo and Hudson recruited all female classes for the study reported in their 2004 paper. In her 1992 doctoral thesis, Ford commented, "Gender, socio-economic status, and ethnic background of the students from each location were considered random."

Twelve studies (11% of the total corpus) analyzed gender in their results, commenting on whether they saw significant differences in their research related to gender. For example, Huppert et al. (2002) reported that “[n]o significant differences were found in the mean scores on the pre-test between boys and girls, within each group and between the groups. These results assure the equal entry behaviour (sic) of both groups.” Other studies found differences in outcomes between genders. For example, Ford (1992) remarks, "The female responses to this series of inquiries resulted in a fairly consistent gender bias in which the consensus was that females were more afraid of snakes than males were."

C. Ethnicity

Only 21 out of 110 papers (19%) made reference to the ethnicity of their study population; the majority (89 papers / 81%) did not. Most of the papers that reported ethnicity did not address ethnicity in the analysis of their data. Only two of the 21 papers reporting on the ethnicity of their study population did more than simply state the ethnic data. In a study looking at the assertiveness of dyad partners and its relationship to conceptual change in students, Windschitl et al. (2001) checked to see if there were any differences in demographics among their three experimental sections. They found that "there were no systematic differences in ability, gender of ethnic distribution across these three sections." Koomen (2006) commented on the ethnicity of the 9 students in her phenomenological study and how it influenced their 'lived space' in the classroom as they learned about monarch butterflies.

D. Other Special Populations

Very few papers - 5 of 110 (4.5%) - address students from other special populations in their studies. The special populations in these 5 papers included students in special education classes, visually impaired students, gifted students and heterogeneous classes.

E. Socioeconomic Status

Fifty-eight of the 110 papers (53%) did not describe the socioeconomic status (SES) of their study population or the urban, suburban or rural setting in which the study took place (Fig. 12). ("Setting" in many cases seems to be a proxy for SES.) The remaining 47% included some mention of SES and/or setting. Ten papers (9%) only mention student or school SES, while 26% (29 papers) reported only on the setting. Thirteen (12%) reported on both. We analyze each of these categories below.

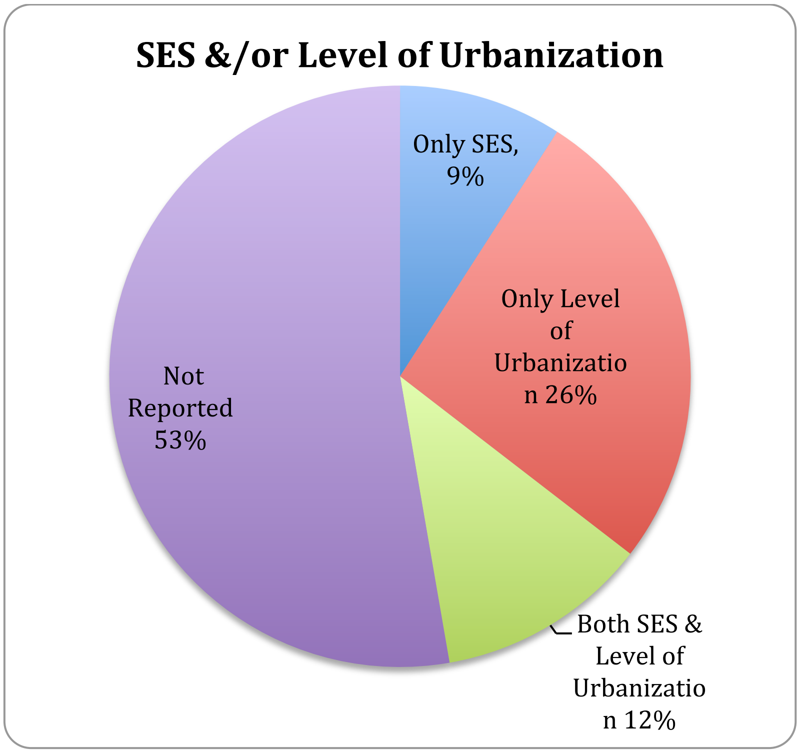


Figure 12.

Three of the 23 papers that described the SES of their study population reported working with a low SES population while 11 papers (48%) reported working with a middle SES population (Fig. 13). No paper reported working with a high SES population. Three of the 23 (13%) reported study populations that included more than one SES category and described these categories in detail (2 papers reported a population that included low, middle and high SES, 1 paper reported low and middle SES). Six of the 23 papers (26%) reported working with populations of mixed SES status, but did not provide any additional information on how these were distributed. Eighty-seven papers did not report on the SES of their study population.

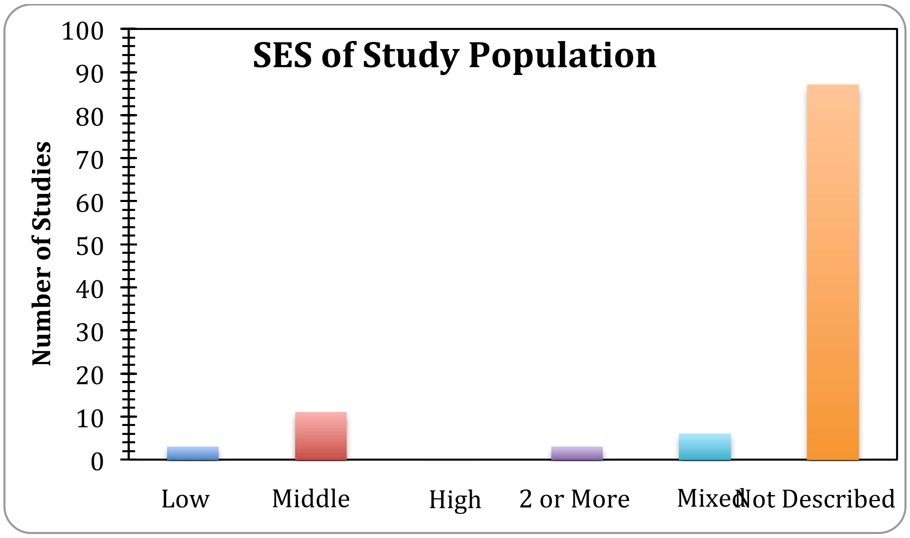


Figure 13.

Twenty-one of the 42 papers reporting on community setting (50%) specified that they worked with an urban population, 5 (12%) reported working with a suburban population and 4 (10%) reported working with solely a rural population (Fig 14). Eleven of the 42 papers (10%) worked with populations from different urban/suburban/rural settings and specifically named these different populations (four worked with mixed urban and suburban populations, four with mixed urban and rural populations, and three with suburban and rural populations). One paper out of 42 (1%) studied a mixed population without specifying the mix. Sixty-eight papers did not report on the urban/suburban/rural setting of the study population.

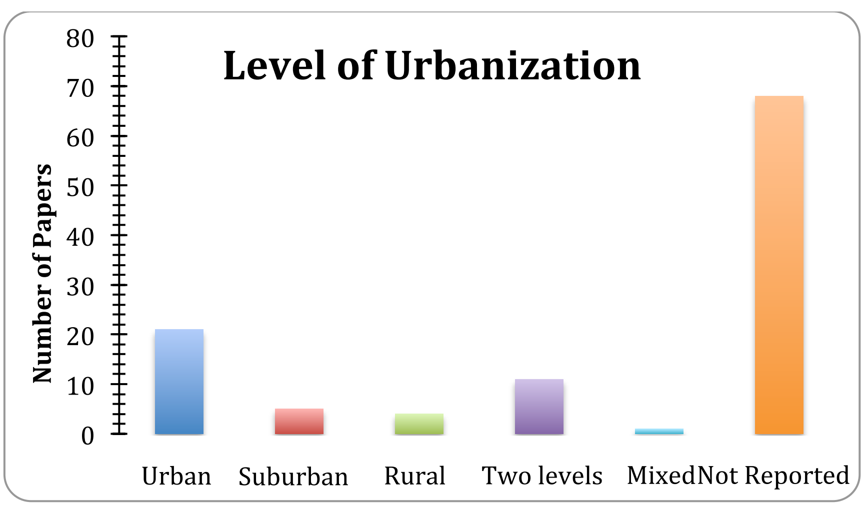


Figure 14.

6. BIOLOGY SUBJECT MATTER

A. General

Ecology proved to be the most common biology domain addressed in the corpus (Fig 15). (36 studies, 32.7%), followed by genetics (24, 21.8%) and natural history (15, 13.6%). In contrast, behavior was mentioned in only one paper. Other topics included Cell Biology, Anatomy, Physiology, Evolution, Biotechnology, and Molecular Biology. Twenty-five papers (23%) reported more than one biology domain; as a result, the number of papers reporting domains is higher than the total number of papers in the corpus.

Eighty-five of the 110 coded papers (77%) focused on only one domain of biology, for example, ecology, cell biology, or genetics. The remaining 25 papers (23%) focused on more than one domain of biology; 16 (14.5%) focused on two biology domains, 3 (2.7%) on three biology domains, 4 (3.6%) on four biology domains, and 2 (1.8%) on more than five biology domains.

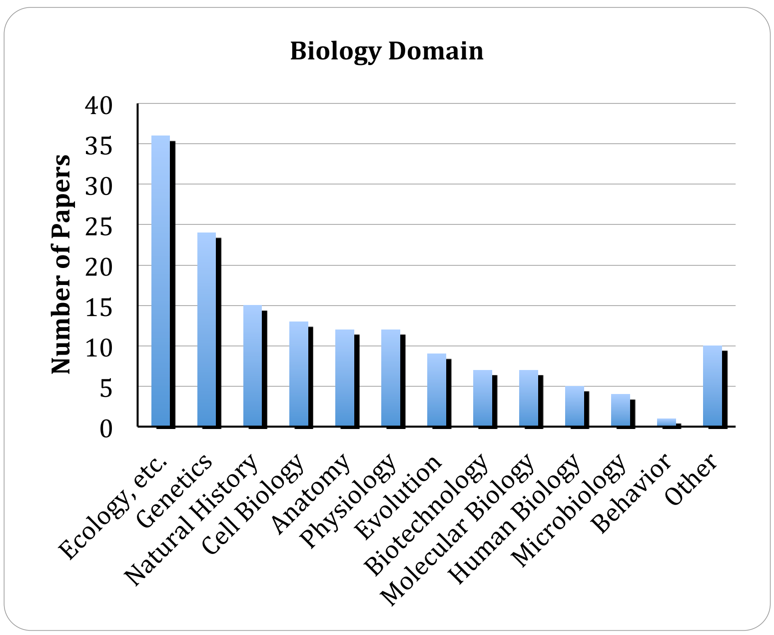


Figure 15.

Grouping studies into broader domain categories reveals an almost even spread across biological scales, with roughly a third of studies in Ecology/Evolution (Fig. 16). A little less than a third in organismal biology and somewhat more than a third focused on the cellular/molecular level. (For the purposes of this analysis, organismal biology includes anatomy, behavior, human biology, natural history; micro/molecular includes biotechnology, cell biology, genetics, molecular biology, physiology, microbiology and virology.)

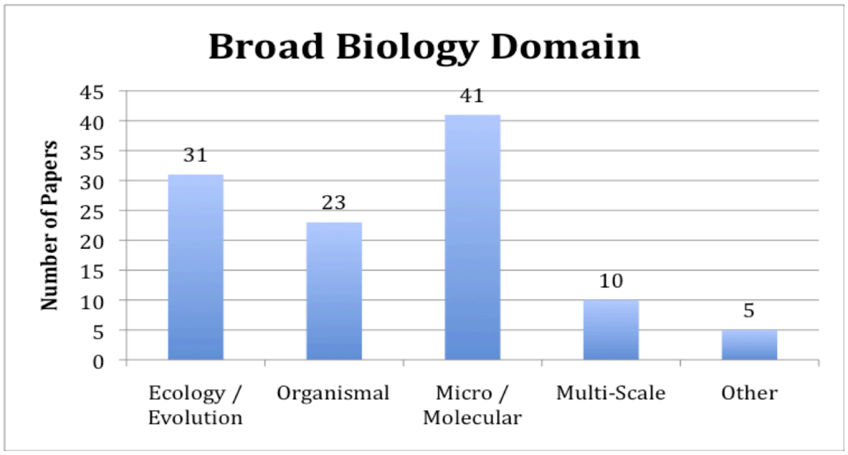


Figure 16.

It is striking that ecology, making about 30% of the total, should be the subject of the largest number of studies. (Note that 5 of the Ecology papers are included in the multi-scale count.) One might conjecture that, given the strong "innovation bias" of the corpus, an increased interested in ecological topics might direct research towards topics in this field. Another possibility is that ecology lends itself to certain innovative approaches that involve systems thinking and modeling, and related reasoning processes. Moreover, ecology lends itself to applied settings that can be used to foster case-based or problem-based learning more easily than some other domains of biology. Finally, ecological topics very often include — or can be designed to include — affective and ethical dimensions which may motivate and engage learners, and lend themselves as well to group work on a challenging problem.

There are studies in our corpus that exemplify all these themes. For example, Hogan (2002) examines "small groups' ecological reasoning while making an environmental management decision." The report provides evidence of the value of case-based problems like this to engage students with a topic and learn relevant concepts, and also evidence that students can reason effectively about complex problems. They note that one group of students, who (according to a pre-lab interview) came in with a more sophisticated understanding of ecology than their peers, consistently showed a richer and more mature and complex approach to the posed situation than their peers: "although this study lends some support to other researchers' claims that adolescents can engage in reasonable discussions about environmental issues even when they have limited prior experience considering those issues, possessing and sharing robust background knowledge does enable groups to analyze, generate, integrate, and evaluate information and ideas to construct a more principled and thorough analysis of a complex environmental problem."

Manzanal (1999) is an example of studies that address the role of ecological experiences — whether laboratory or simulation based, or fieldwork based, on student attitudes. In the case of this study, the focus is on students' attitudes towards environmental protection. Randler and Hulde (2007) show that hands-on activities relating to soil science with middle-school students resulted in more student learning, and more engagement, than a lecture-demo approach, and that these effects were discernible when tested 4 weeks after the unit was concluded. Laflamme's (2004) dissertation engaged people from several age groups — elementary school to young adult — in a process of learning about a local aquatic environment, and about fish from the area, both in their habitat and in aquarium and other settings. He found that as the participants grew in their familiarity with the fish, they also grew in their capacity for empathy for these organisms so very different from themselves.

For the papers dealing with the other topic domains, conceptual understanding, or disciplinary reasoning, are the primary goals. In evolutionary biology and in genetics, computer environments of various kinds provide students with an opportunity to explore challenging conceptual fields, supported by structured, interactive programs that typically allow repetition or replication of procedures. The systems do not have the richness and complexity of living systems, but provide learning environments that are engaging and scaffold some opportunities for student inquiry, data analysis, and scientific reasoning. Some such environments can include material that is "beyond grade level," so that the technology is accessible to students of different levels of attainment, understanding, or interest.

Koomen (2006) describes a hands-on activity, focused on the rearing and study of Monarch butterflies, in which one aspect of living systems plays an important role in the content of the classroom. While the teacher was using a published curriculum on monarch biology, she had to improvise when a disease struck the class's butterflies, and many of them died. This required the teacher to shift the focus of the lessons, though the study does not report how the shift of focus to a somewhat more abstract theme — insect ecology — affected students' learning or engagement. As in LaFlamme 2006, the involvement with living organisms can increase students' attentiveness and investment in a lab experience, but can also be aversive: people have attitudes about different organisms, and insects have for some students a certain amount of "ick."

One of the drawbacks of experiential lessons is that it can be difficult to know what is to be learned, what the patterns are which the students are to see, what conclusions can be drawn from the evidence. This means that it lies more with the teacher to help students draw definite conclusions (perhaps individually), which can help drive learning forward, and thus the pedagogy associated with the lab activity or learning environment plays an important, perhaps decisive, role. In the monarch unit described in Koomen 2006, the observations of butterflies (at various life stages) were supplemented by teacher presentations, and by games that helped students understand and play with concepts and relationships, which for some helped clarify what they were seeing in their observations of live creatures. This and other studies thus reflect the established insight that possible topics need to be approached in more than one way. Living organisms add complexity, surprise, and in some cases vividness.

B. Nature and Process of Science

Just eleven papers (10%) addressed student learning about the nature and process of science as part of their research in addition to a biology domain (Fig 17). Three papers addressed the nature of science, 7 papers address the process of science, and 1 paper addressed both the nature and process of science. The sample was too small to look for possible relationships to particular biology domains.

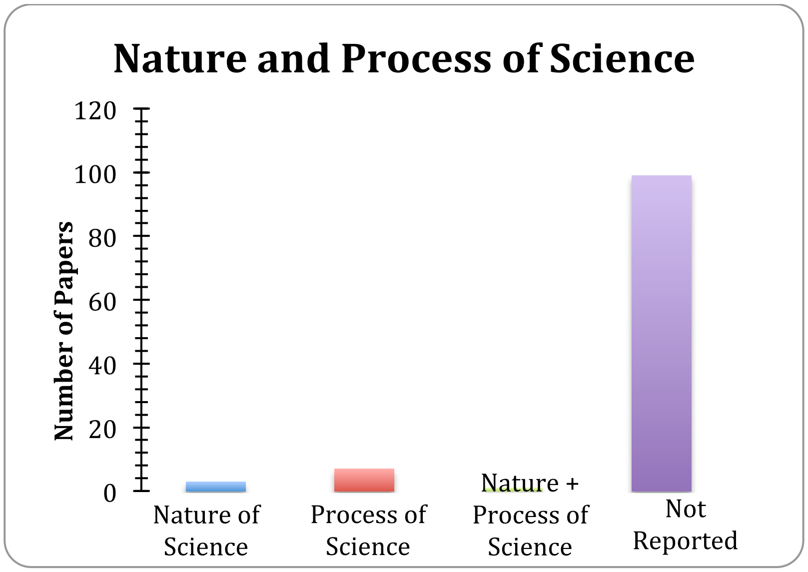


Figure 17.

7. PEDAGOGICAL DESIGN

Under this heading we include such characteristics as the instructional purpose, its relationship to the overall curriculum, and the "degree of inquiry" in the pedagogical framework for a laboratory experience.

A. Types of labs: materials and methods used

What type of activity or technique was studied? Students engaged with whole live organisms in just over a quarter (27%) of the studies; a further 40% used prepared biological material (e.g., tissue samples, slide preparations) (Fig. 18). Modeling of some description was the most frequent type of lab students engaged in; over half (52%) involved the use of physical or computerized models often supported by other multimedia materials. Only 14% described student engagement in quantitative activities, such as data analysis, while just over a fifth (22%) included print materials. Note that 48 papers (44%) reported that students studied only one type of lab or material, while the rest included two or more types of labs or materials.

Computer simulations were used for instruction in many studies involving genetics, molecular biology and cell biology on the one hand (55%), and ecology and evolution (36%) on the other. For the most part, students engaged with simulations for extended periods of time, and authors for the most part describe how students' work was carefully scaffolded to deal with complex and abstract phenomena or to help students engage with invisible phenomena. Providing an opportunity for students to engage with complex data was an explicit feature of many of the computer simulations, since they can support students to make inferences and justify them.

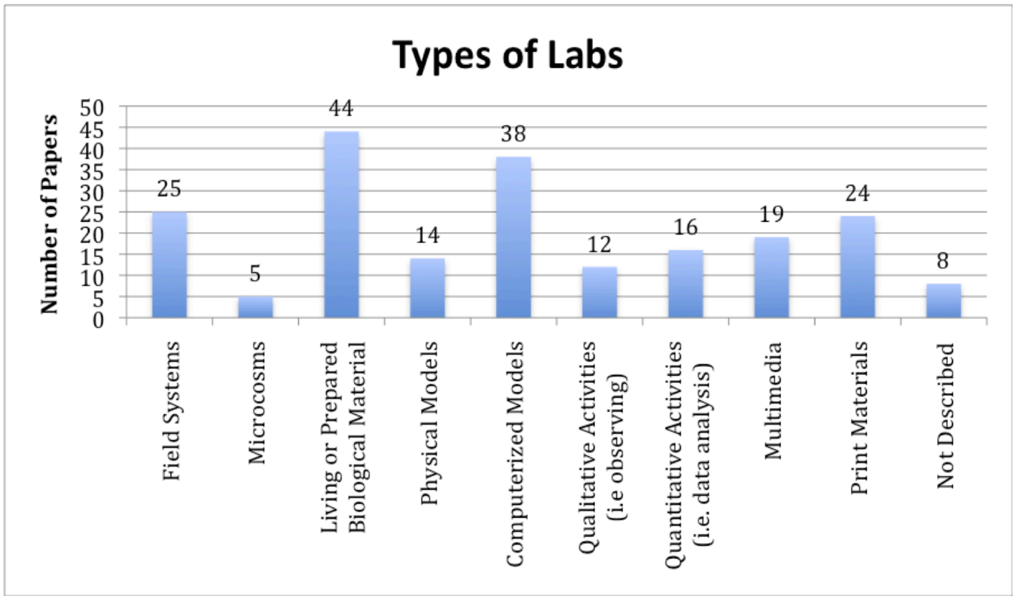


Figure 18.

B. Instructional purpose

What was the author's stated instructional purpose for the activity or technique? A large majority of papers (98 or 89%) reported learning new concepts as one of the instructional purposes of the laboratory activity studied (Fig. 19). 'Demonstration' was the category least often studied (3 papers). Other instructional purposes were 'manipulative or experimental technique' (39 papers), 'analytic technique' (50 papers), 'exploratory hands-on activity' (15 papers).

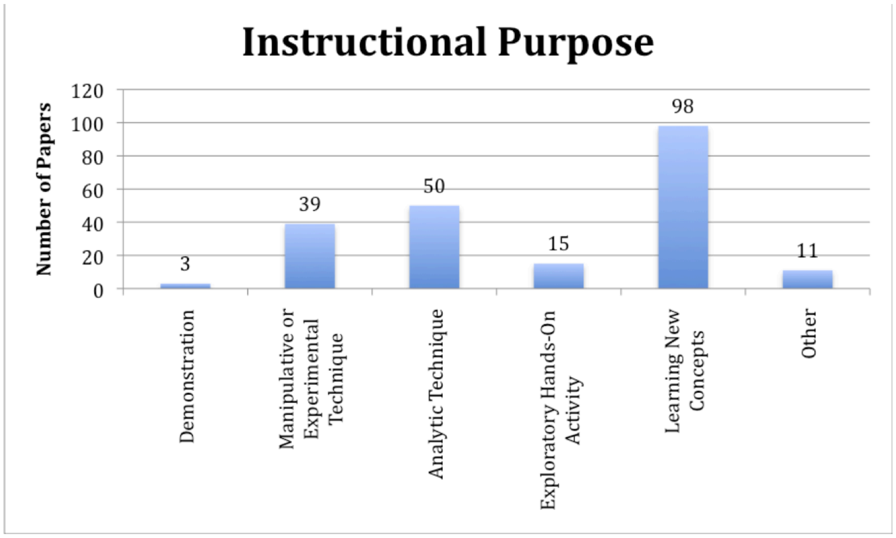


Figure 19.

What sorts of lab activities did authors study as ways to meet the instructional purposes described? Students engaged in observing, exploring and gathering observational data in a large majority (94 or 89%) of the studies (Fig. 20). They engaged with lab tools and procedures, including the use of models, in all of the studies, analyzed data in over half (57%) of them, and engaged in argumentation in under half (40%). A wonderful 15% involved students in taxonomic classification of some sort (wonderful given the much lamented dearth of students interested in taxonomy at college levels and higher-perhaps there is hope after all). Finally, almost a quarter (24%) reported that the lab activity focused on conducting full experimental investigations. Note that, not surprisingly, the large majority of papers (86%) described labs in which students were engaged in more than 1 activity; 15 described only one activity type.

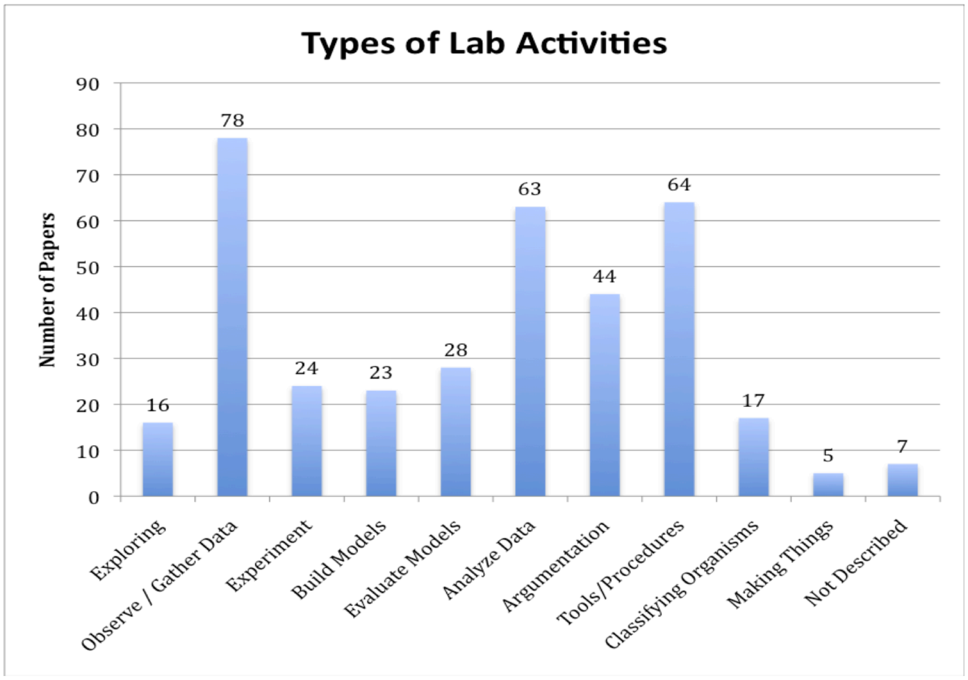


Figure 20.

C. Length of Lab Activity

Of the 110 coded papers, 45 papers (41%) reported that the lab activities on which they conducted research only lasted the length of one class period (or less). Sixty-five papers (59%) reported studying lab activities that lasted two or more class periods.

D. Degree of inquiry

To define inquiry, we measure such characteristics as the amount of scaffolding provided, either by teacher or curriculum materials, with respect to relative importance of student and teacher responsibility for (i) question selection, (ii) the adoption/critique of methods, and data analysis, and (iii) sense-making and relation to explanatory theory. What are some of the contexts in which aspects of inquiry do make their appearance in the education research literature at large? What motivates choice of domain when inquiry is deployed as an instructional strategy?

In addition to the variety in types of science investigations represented in laboratory exercises, we also inquired into the pedagogical setting that mediates the way students engage in these activities: What kind of intellectual work are the students doing, or should they be doing (Drayton and Falk 2001, Monk and Osborne 2000, Millar et al. 2000)? What kinds of argumentation and reasoning are expected of them (Driver et al 2000)?

It is difficult to tell exactly what classroom intervention students are actually engaged in from an investigator's description in a publication, as Minner et al (2010) and others have noted (and see "Gap Analysis" below.) This was the case in our study also. For example, in the following composite passage, typical of many we read, it is difficult to tell if the students took responsibility for their own learning and were engaged in inquiry: "…students engaged with "cases" posed as problems, that they solved by making predictions, testing their predictions by selecting relevant data from the dataset, analyzing data, and stating their conclusions."

Authors describe activity structures, but in this composite we cannot discern the range of possible predictions. For example: Were possible predictions few or many? How much choice did students have when selecting data? Were data analysis methods prescribed? Did teachers ask student to provide warrants for their conclusions?

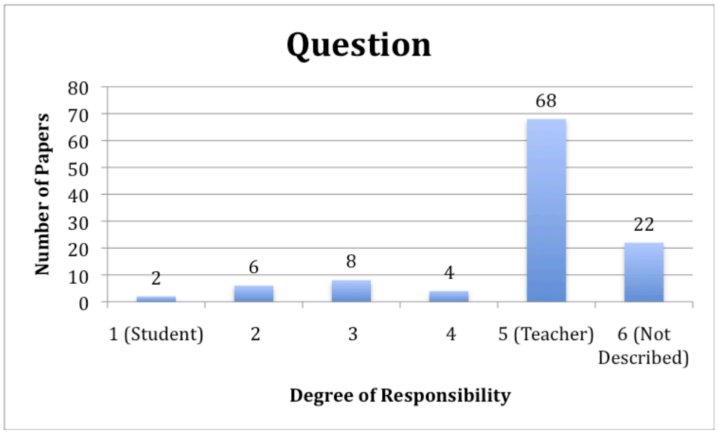
In many cases, we had slim evidence to draw on, for example, often needing to take a statement from the author that students were engaged by means of "constructivist methods" as evidence that we might be in the presence of inquiry. We obviously required more evidence that this, so we defined degree of inquiry as the extent to which students were responsible for a) the **question** being investigated in the lab, b) the **design** of the investigation, and c) the **construction** of knowledge. We assumed that these three core components of a lab activity provide a comprehensive enough assessment of the degree to which students were responsible for their own learning through inquiry. A code of 1 was assigned if the student was completely responsible for each of the three components of inquiry, while a code of 5 meant that the teacher or the curriculum was responsible. If it was possible to determine that some degree of inquiry was incorporated into the design of the lab in question, a default code of 3 (in a scale of 1 to 5) was assigned (Table 1).

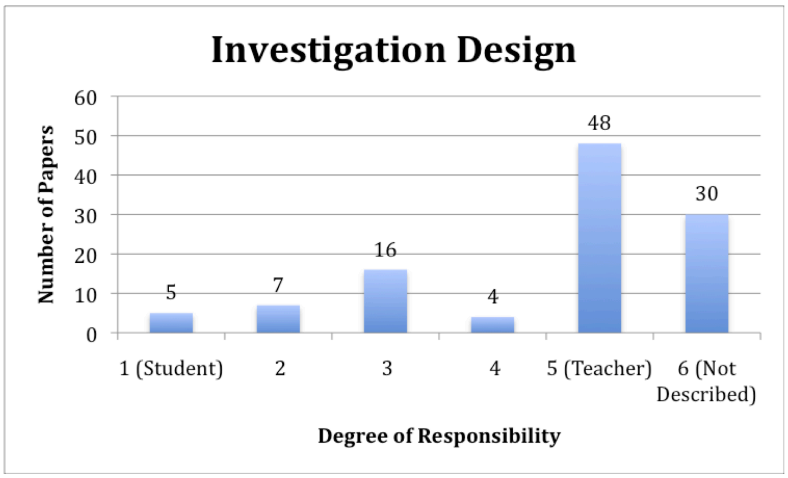
Degree of Inquiry					
24. Degree of responsibility for Question					
1___ (student)	2___	3___	4___	5___ (teacher)	6___Not described
25. Degree of responsibility for Investigation design incl. data collection					
1___ (student)	2___	3___	4___	5___ (teacher)	6___Not described
27. Degree of responsibility for Outcome and Construction of knowledge					
1___ (student)	2___	3___	4___	5___ (teacher)	6___Not described

Table 1. Degree of Inquiry Codes

We focus the discussion reported here on "inquiry scores" of 3 or less, since a code of 3 represents the midpoint between student and teacher responsibility for students' inquiry activities.

Fourteen percent of the 110 papers in the corpus (16 papers) had an inquiry score of 3 or less posing their own investigation questions, 26% (29 articles) for design of the investigation, and 50% (55 articles) for making sense of the outcomes of their investigations and constructing their own knowledge (Figs. 21a, b, c).





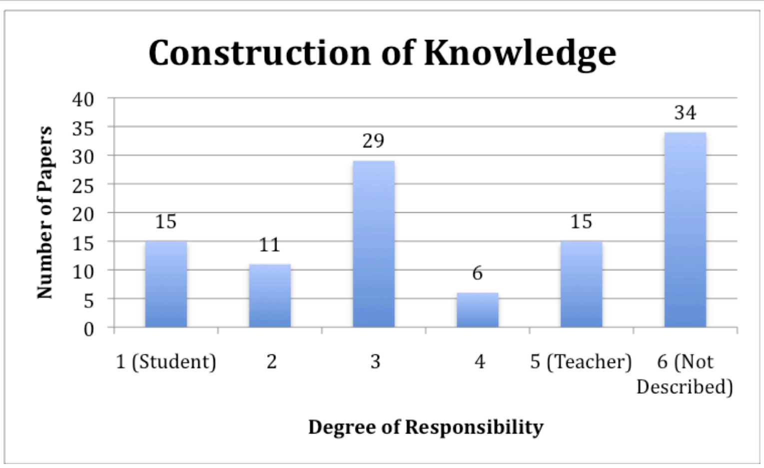


Figure 21a, b, c.

In aggregate, just over half of the corpus (59 papers) had an inquiry score of 3 or less for at least one of the three inquiry measures. Minner et al. (2010) describe inclusion/exclusion criteria that defined articles as "inquiry-based" if they instructed [students] via some part of the investigation cycle (question, design, data, conclusion, communication)" and included "pedagogical practices that emphasized to some extent student responsibility for learning" (p. 479). Following these criteria, we can conclude that a good deal of lab-based instruction in life sciences is conducted using inquiry-based methods, though this gives us no information about the inquiry orientation of the course in which the labs were (sometimes) embedded.

While students may not always be free to decide which questions to investigate, they are taking some responsibility for asking questions. Further, almost half of the lab experiences described in the "inquiry corpus" appear to offer the possibility of inculcating general understanding of the nature of science and scientific practice. Students are taking some responsibility for designing their own investigations, for example, choosing and/or identifying appropriate variables, setting appropriate levels for quantitative variables in computer simulations that address their predictions, and considering whether their designs will result in needed data.

Responsibility for constructing their own knowledge was the inquiry measure that included the highest number (55) of articles, while in the other half the outcome of the lab was predetermined by the teacher or the curriculum for the lab in question. Descriptions of student construction of knowledge included student responsibility for strategies to organize their data and make sense of it, construction of conclusions and ensuring that their conclusions were supported by their data, and relating findings to prior knowledge or to scientific principles. Note the overlap that many of these items have with the student reasoning outcomes described later.

It is not surprising that the great majority of the studies focused on 8^{th} grade or higher, given that the great majority of this body of articles, namely 37 of the 59, focused on invisible, complex, or abstract phenomena, related in many cases to the biological distinctives. For example, studies in ecology or evolution focused on complex ecological modeling (Stratford, Krajcik and Soloway, 1998), thinking about complex systems and understanding how ecologists construct and use dynamic models as tools to develop insights on the relationship between system structure and function (Hogan and Thomas, 2001), and "reasoning [in genetics] from effects (phenotype) to causes (genotype)" (Slack and Stewart, 1990).

In addition, the majority of the labs that incorporate inquiry in this more rigorous sense extend, not surprisingly, over several class periods. Twenty-two of the 59 labs (37%) occurred during 1 class period, 15 occupied 2-4 periods and the remaining 22 ranged from 1 week to semester-long (Fig. 22)

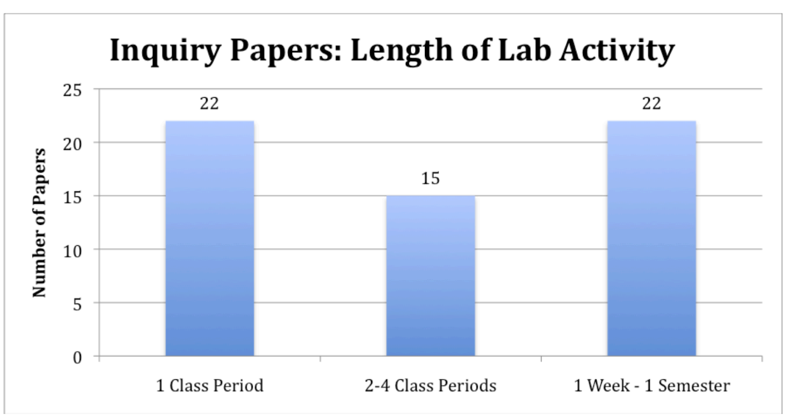


Figure 22.

Extended engagement with phenomena, as well as giving students the opportunity to reason, are the stated motivators of many studies in this subset of the corpus. Frequently, in setting the context for their studies, authors describe one goal of their instructional interventions as providing an opportunity for students to think like scientists do, to reason, to make sense. For example, Hafner and Stewart (1995) state: "Situating conceptual knowledge of a discipline in the context of its use in the solving of problems allows students the opportunity to develop…insights into the nature of the discipline [genetics] as intellectual activity" (p. 111). Authors describe "reasoning" in several different ways, as promoting "causal reasoning skills" (Zohar, 1996), engaging in "reflective thinking" (Seethaler and Linn, 2001), or changing the way students view scientific knowledge in which "the meaning of data is debated and theories are not absolute" (Taraban et al., 2007).

Can a lab experience be considered constructivist if, in spite of a predetermined outcome, students have responsibility to construct their understanding? It seems that the degree of engagement and freedom experienced by the student might hinge on the way the lab is framed either by the curricular materials or the teacher. Opportunities to construct one's own knowledge can be carefully provided in a lab setting in order to support autonomy and to counter the student request prevalent in high school classrooms for "the right answer." We conjecture that, if the system being investigated is complex and dynamic, highly specific instruction could still yield emergent or undetermined knowledge construction for students even if the lab specifies a predetermined outcome. However, we have yet to examine the remaining half of the corpus that was not "inquiry-based" to address this question.

E. Student data analysis

Seventy-five papers (68%) reported that students analyzed data from only one source, 31(31%) reported two different sources, 3 (3%) from three different sources, and one from four different sources. What types of data did students analyze? Thirty-five of the 110 papers reported that students analyzed data from more than one source. Fifty-nine papers reported that students analyzed qualitative data. In forty of these, students collected the data themselves, while data was provided in the remaining nineteen. Forty-five papers reported that students analyzed quantitative data. In thirty-four of these, students collected the data themselves, while data was provided in the remaining eleven. Forty-six papers did not describe student data analysis (Fig. 23).

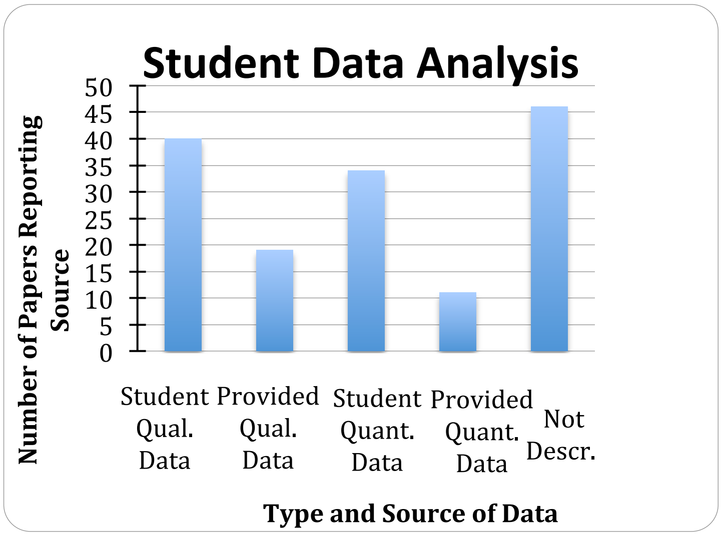


Figure 23.

8. PROGRESSIONS

The uneven coverage of subject domains means that the papers themselves do not provide the basis for proposing any progression of skills or topics in the laboratory curriculum of life science, though many of the papers designed their study to address prerequisites for learning the specific subject matter in the study, and thus furnished possible hypotheses for further study on progressions within, say, evolution, genetics, or ecology. The majority (77% or 85 papers) of the 110 coded papers reported that the lab activity of the study was related to a broader biological or pedagogical context, that is, was designed to improve student outcomes on a point explicitly related to the learning of some later or more complex topic. None, however, examined whether positive student outcomes in the study translated into later improvements in learning. Seven (6%) did not connect the lab activity to a broader biological or pedagogical context. Fifteen percent (17 papers) did not include information on the context of the lab. However, only one paper (1%) specifically studied a laboratory experience that was linked to a progression across grades. Thus, in addition to uneven coverage, another reason the corpus makes little contribution to the identification and validation of learning progressions is that almost none of the studies were (at least explicitly) embedded in a larger-scale research program to accomplish that purpose.

9. BIOLOGICAL DISTINCTIVES

A. Overview

A key part of the rationale for Under the Microscope stems from the concern that, despite the critical importance of life sciences (including ecology and public health) for social welfare and personal interest, the evidence is that [a] the public in general does not learn biology very well,[2] [b] only a minority of students find their life science courses interesting, and [c] there is a serious shortage of people going into several key areas of biological science. (NRC, 1990).

Various causes have been identified. The NRC study condemns the science that students encounter in broad terms: 'Previous exposure to science, minimal as it is, has burdened the subject with mystique. Instead of being seen as the way to infer relationships and causes through observation and trial (experiment), which most people engage in to various extents in other parts of their lives without thinking about what they are doing as "science," science is viewed as arcane, difficult, practiced only by the very talented, and unrelated to the real world of the average person. For most students, instead of dispelling those notions, the tenth-grade biology course simply reinforces them.' (pg 10). The authors lay the blame on textbooks (too long, too detailed, too boring), on teachers (who rely heavily on the long, boring, badly written texts), and on university instructors who do nothing to dispel these various discouragements.

Many years ago, Choppin and Frankel (1976) found that more than half the lessons reported by students as their 'peak learning experiences' in biology were in laboratory sessions, when the students themselves conducted experiments. Nevertheless, biology lessons in the secondary school typically begin with the teacher explaining what the class is going to do. Students have no choice of topics and the teacher rarely uses their ideas. In the laboratory, students usually work on problems set by the teacher, and only sometimes work out their own method (Trumper, 2006).

The National Research Council (1990) has some suggestions for directions to make it better. They see Natural History as an important focus for elementary school, human systems for middle school, and then some meaningful mix of cell and molecular biology, energy and metabolism, ecology, and evolution for high school.

B. The distinctives introduced

Reading biology texts, and the literature on student interest, and the nature of modern biology, we were drawn to conclude that one result of the current mainstream approach in biology education is that the density and formality of the presentation actually prevents students from encountering aspects of the living world that are engaging to the mind and imagination. In addition, biology texts and courses tend to reflect primarily a modern view of the organization of biological knowledge, which places a high emphasis on the current understandings of fundamental mechanisms, which is not necessarily the best way to get into a subject, or understand its interest and implications. A shorthand way of summarizing the results is that biology students have remarkably little contact with living systems.[3]

---

[2] The NRC's Committee on Life Sciences remarks (NRC 1990, page 12): "By any reasonable measure, most high-school students graduate without knowing even the rudiments of basic biological concepts. The students therefore leave school with deep misconceptions about biology that may seriously affect their lives."

[3] Dewey articulates this problem very incisively in Democracy and Education (Ch. 17): "Logical order is not a form imposed upon what is known; it is the proper form of knowledge as perfected. For it means that the statement of subject matter is of a nature to exhibit to one who understands it the premises from which

Furthermore, the typical course does little to engage students with the tremendous diversity of biological studies, as reflected in the many fields of biology. While the goal of the K-12 system is not to produce more biologists, one result of the students' encounter with life science classes during their school career should be that they understand something of the importance and interest of the field(s), and the very diversity of biology can provide an entry to such an understanding for students of quite diverse interests and backgrounds.

In our proposal, we argued for a program of research and development in life science education that should "take explicit account of features that are unique or highly characteristic of biological systems, and therefore of the reasoning and methodologies that are appropriate to these distinctive features. " But what aspects of these systems might be desirable to include, so as to change this state of affairs? The next section reproduces the presentation in our proposal. We follow this with some reflections on each of the distinctives in more detail, as it relates to our coding scheme, to the "Microscope" study, and to our larger program.

C. What sets biology apart?

> *There are no inanimate systems in the mesocosmos that are even anywhere near as complex as the biological systems...These systems are rich in emergent properties...are open systems - richly endowed with capacities such as reproduction, metabolism, replication, regulation, adaptedness, growth, and hierarchical organization. Nothing of the sort exists in the inanimate world* (Mayr 2004, p. 29).

It is rare for researchers or curriculum developers in science education to address the differences among the sciences explicitly (Bell 2004). There are good reasons for this,

---

it follows and the conclusions to which it points ...To the non-expert, however, this perfected form is a stumbling block. Just because the material is stated with reference to the furtherance of knowledge as an end in itself, its connections with the material of everyday life are hidden. To the layman the bones are a mere curiosity. Until he had mastered the principles of zoology, his efforts to make anything out of them would be random and blind. From the standpoint of the learner scientific form is an ideal to be achieved, not a starting point from which to set out. It is, nevertheless, a frequent practice to start in instruction with the rudiments of science somewhat simplified. The necessary consequence is an isolation of science from significant experience. The pupil learns symbols without the key to their meaning. He acquires a technical body of information without ability to trace its connections with the objects and operations with which he is familiar -- often he acquires simply a peculiar vocabulary.

There is a strong temptation to assume that presenting subject matter in its perfected form provides a royal road to learning. What more natural than to suppose that the immature can be saved time and energy, and be protected from needless error by commencing where competent inquirers have left off? The outcome is written large in the history of education. Pupils begin their study of science with texts in which the subject is organized into topics according to the order of the specialist. Technical concepts, with their definitions, are introduced at the outset. Laws are introduced at a very early stage, with at best a few indications of the way in which they were arrived at. The pupils learn a "science" instead of learning the scientific way of treating the familiar material of ordinary experience. The method of the advanced student dominates college teaching; the approach of the college is transferred into the high school, and so down the line, with such omissions as may make the subject easier."

since the sciences have in common a wide range of values and techniques, and "scientific reasoning" can usefully be characterized in generic terms. However, it is also important to recognize that, in fact, there are many sciences, reflecting the diversity of the world and of human questions and interests. An effective science education will strengthen students' understanding of the world and of the several kinds of science, by taking these key differences into account, as well as clearly affirming the core of values and practices that all the sciences share.

Several key characteristics taken together differentiate biology from other sciences (Mayr 2004, 2002, Sterelny and Griffiths 1999). These include:

- The great complexity of living systems, the ways they are organized and inter-related, and the properties that emerge from these complex systems at different levels of organization
- Evolution, which plays a unique, cohesive role across all domains of biology, and often requires historical and comparative reasoning
- Populational thinking, and the significance of individual variation
- The construction of biological theories around concepts rather than laws
- Dual causation: Causation of biological phenomena include interacting genetic and environmental contributions.

The distinctive characteristics of biology have implications for the relative importance of observational, experimental, and comparative studies in the biological sciences. They shape the sort of representations (quantitative and qualitative) that are useful and powerful in biological research, and the kinds of inferences and certainties that are possible (e.g., Stratford et al. 1998, Rudolph and Stewart 1998, Passmore and Stewart 2002). Beyond such abstract factors, however, there is the over-arching fact that biology deals with living systems, and these impose their own constraints and limitations on the investigator. Growth in biological understanding requires significant, scaffolded engagement with the complexity and variation in living systems, the challenge of recognizing signal and noise, and the unpredictability, contingent, or emergent nature of many biological phenomena.

Not all of these considerations should be presented in the same way for, say, a 6^{th} grade life sciences class looking at the idea of food webs, a 10^{th} grade class doing genetic experiments with Fast Plants, or college students studying *Drosophila* embryology. However, we suggest that the failure to properly incorporate them into the presentation of biological ideas, and into students' experiences with biological investigations, has led to the common perception among students that biology is really an endless list of facts and arcane terms, studied by disparate and apparently unrelated fields of science. The underlying coherence of the biological view of life, and the complexity and diversity of life, must be conveyed from the beginning of students' encounter with the life sciences (Janovy 2003).

Finally, many of these characteristics are accessible at various levels of qualitative and quantitative rigor. As a result, we might expect that students will gain a more and more mature appreciation of these characteristics over the course of a life-science learning career as they learn more and more about the living world (including our own species), and as they gain in their ability to investigate natural phenomena, and reason about them (Driver et al. 1996). As the only curricular locus in which these characteristics are presented, one can expect that biological laboratory experiences should reflect them. Furthermore, we might expect that research on biological lab experiences would address these topics in some way, and provide insight about whether they improve students' understanding of living systems. We might even expect that laboratory experiences will increase in their demands on students' understanding of these aspects of living systems as the students advance through their life-science education.

D. Unpacking the distinctives

In this section, we discuss each of the distinctives in turn, in the context of our study.

1.Complexity. We have agreed, in discussing the coding scheme, that a key aspect of this concept is the notion of levels of organization. Any biological phenomenon is embedded in a hierarchy of organization, from molecule to ecosystem. But this truism, while true, has implications that may not be evident at first. An important related fact is that each phenomenon is embedded in a web of relationships or interactions with its environment. For a biomolecule, the environment typically is the cell within which it operates, and more especially the other molecules with which it interacts, according to the rules of chemistry and physics. The codons of a DNA molecule interact with other elements of the genome by way of various enzymes, RNA factors, etc., and these in turn are "motivated" by chemical and physical conditions; messengers run in both directions, to and from the genes. Furthermore, these interactions cross levels of organization: cells to tissues, tissues to organism, etc.

The complexity of biological systems is related to the phenomenon of emergence: properties that are present at higher levels of organization but not below them, thus at the cellular level but not the molecular, at the organismal level but not the cellular level, and so forth. While in many cases these properties can be explained satisfactorily in reference to lower-level mechanisms, in other cases this is either not (yet) possible or meaningless. While to a certain degree one can say that oxytocin promotes attachment between members of a social group, the experience and phenomena of attachment are not completely describable in terms of hormonal levels and actions.

Even a process like osmosis, which is often exemplified in labs (sometimes in relation to the behavior of membranes), is "complex" in part because of the conditions which modulate its operation, and in part because of the many, many situations in which it plays a role in living systems.

Now, how might this kind of complexity be addressed in the lab course? We have outlined in the codebook several loci for the distinctives:

a. The author addresses the "distinctive" in the intro, theoretical framework, results, or discussion

b. The students address it in prep for the lab

c. The students address it in the lab itself

d. The students address it in the sense-making part of the lab.

From the point of view of our current, limited study, the principal question is, Does the research examine whether the labs help students understand the complexity of the phenomenon — or is the phenomenon examined only in isolation? Extending this, however, by reflection on the 4 loci for which we are coding, we can ask, for example, whether the authors understand or articulate the complex biological implications or connections of the lab's content (this content takes additional importance from the links to other body systems, organisms. etc.), or the relationship of the lab's content to the students' understanding of more than the specific phenomenon. Or are the students asked or encouraged to address such questions? Or, perhaps even more interesting, do they have to make sense of the lab's requirements or outcomes in the context of (some of its) interactions?

Let us examine two papers from our corpus, (Akpan and Andre, 1999, a lab on frog dissections virtual and real) and (Ergazaki and Zogza, 2008, a lab on lake ecology using a simulation). The first we coded as not dealing with complexity, the second did. Ergazaki and Zogza address complexity in more than one way. The activity requires the students to discuss data created by a simulation that address fish populations and water quality, among other issues. The exercise is explicitly multilevel; some dyads in the class actually are aware of that, and then the authors in the "implications" talk about the importance of thinking about population dynamics, in understanding regulation of populations and the idea of "equilibrium" in terms of resources etc.

Akpan and Andre, on the other hand, focus entirely on the challenge of identifying and extracting frog parts (though one of the objectives of the lab is to "learn the functions of the frog skin for protection from predation through camouflage and secretion of poison"), and learning the functions of each frog part identified and extracted. What more should be expected? We inspected two prominent biology texts (Miller and Levine, Starr and Taggart) to see how they dealt with organs — presumably the units alongside which the dissections might occur. In both cases, the organs were presented very effectively in the context both of the systems they fit into, and their physiology. Thus, to really accompany such texts, one would not necessarily dissect all available frog parts, but explore kidneys, nerves, etc., in the context of their systems' functioning and regulation. This places the organs in a systemic context, and in fact Miller and Levine, in particular, take opportunities to note evolutionary dimensions as well.

An alternative approach that could present the material in a coherent fashion while also preserving some of the complexity implied in the topic would be the framework of comparative anatomy, or even historical/evolutionary development. Obviously, organs for pumping blood could be compared and contrasted, and within the vertebrates could also be arranged in historical sequence, and connected (perhaps) with eco-physiological considerations.

The point is not that Akpan and Andre should have done something different, but rather that different approaches might've been taken, which could implement the traditional goals of a dissection in a conceptual context more informed by the approaches of a well-designed textual narrative, which militates against the presentation of biological information as a collection of isolated facts.

<u>2. Dual causation</u>. It is important for students to understand that biological activity is "information driven," in the sense that most physiological events (and such events underlie essentially all biological activity) are initiated or regulated by the expression of

genes. However, organisms are also influenced by their environment, during growth and development as well as throughout life. This way of thinking about causality embraces the complex idea of the relationship between genotype and phenotype, the notion of evolutionary constraints, and the role of basic principles of physics and chemistry.

Obviously, the genetic contribution to reproduction, homeostasis, and behavior is complex, and there is a lot to learn in order to understand the mechanisms and dynamics of gene expression, and the relation of genotype and phenotype. Typically, while genotype/phenotype relationships are addressed in many labs reported in our study (23 papers), there is rarely any evidence that the labs present this relationship as involving any complications: the specific genome reliably produces the trait in question.

A next layer of complexity includes environmental influences upon gene *expression*. A common lab that addresses this explores the effects of temperature upon the rate of chemical reactions and enzymatic activity.

Another layer, however, is little explored in lab activities: biophysics. This fascinating area is sometimes touched on in relation to plant physiology (e.g. evapo-transpiration), and occasionally (as in the AP Labs) in other topics (circulatory system), but otherwise it is neglected. The behavioral, physiological, ecological, and evolutionary facets of the physical constraints upon living systems are full of possible investigations and activities, and could provide a powerful setting for exploring "core" topics in the curriculum.

<u>3. Evolution</u>. This "distinctive" may show up in many different forms — from lessons about comparative anatomy, to transmission genetics, to natural selection, to biodiversity and speciation. Many texts address briefly the history leading up to Darwin's work, and then briefly the rise of the "modern synthesis," and even some controversies such as punk eek. A "tree of life" is often depicted. The Hardy-Weinberg equilibrium is introduced. Classification and the history of life are also typically included in the texts, and biogeography as a source of evidence for evolution. Miller and Levine mention HOX genes and genetic taxonomy, including a brief introduction to cladistics (sparing us synapomorphies and similar verbal monstrosities). Evo-devo (evolutionary developmental biology) is mostly not mentioned, except in reference to embryology.

Another aspect of this topic that is little addressed, and yet critical for understanding the mechanisms at work, is the question, "On what does natural selection operate?" This is not only about the "unit of selection," but also about which traits are the direct point of "scrutiny" by natural selection, nor merely the collateral consequences of it. What is an adaptation vs. a contingent trait?

One might hope that the evolutionary dimension of most topics should be addressed, at least as a comment which explicitly reminds the reader that anything we look at in the biological world is an evolutionary result: We've been talking about reproduction or respiration in a generalized eukaryotic cell. What are some of the varieties found elsewhere in the living world, and what can we say about the developments over time?

<u>4. Populational thinking</u>. This is a foundation stone of evolutionary biology. In the first instance, it emphasizes that there is no "type" of a species, of which individuals are imperfect instantiations (there is an interesting echo of typological thinking in object-

oriented programming, in which a kind-of entity is defined by a data structure and associated “methods”, but the actual computation is performed using one or more instantiations of this pre-defined class of objects). Rather, individuals exist in populations, and populations are located in actual conditions in time and space. The biological species concept provides a fairly straightforward way to delineate populations, for some taxa.

The various levels of biological classification must be reconceptualized from this point of view, and the ontological status of each has come under question. To what extent is a species a thing in nature, as opposed to a human construction? A class? A kingdom? These notions all must be relatable in some realistic and consistent way to populations (though some relief has been afforded by the development of the idea of metapopulations).

But an additional implication of this point of view (or as Mayr calls it, the theory of population thinking) is that every organism (and probably every cell in every organism) is unique. Some of the differences may escape our ability to detect them, but it is true that each organic system has both an inheritance and a personal history: “nature” and “nurture,” if you like, or genetics and environment. What we see, mostly, is the result of the two sources of characteristics; it requires further analysis to differentiate them.

Nevertheless, the variety in some sense *is* the population (and by extension, the species) — it is not just values around the mean, or within one or two standard deviations, it is the whole distribution that defines the group. The search for the sources of variations is at the heart of a lot of biology, and so each individual is a genetic, an ecological, and a historical entity.

There is something else here which is not often mentioned in this connection, in introductory biology. This is that we do not vary with respect to one trait at a time. Variation in size and shape, reaction times and intelligence, co-occur with variations in location and size of organs, tolerance for various toxins, etc. etc. Therefore, it is very rare that selection “sees” a specific trait and applies pressure on it. It is not surprising that in making discriminations about mates, for example, that organisms may use a proxy to identify superiority, and that this proxy may provide indirect evidence about a bundle of traits, some at least of which may be heritable. A female manikin or bowerbird who chooses her mate on the basis of dancing skill or bower beauty is selecting a mate whose attractiveness is rooted in a lot of elements.

The isolation of one trait of interest, and the ignoring of the elemental truth that traits are specific characteristics of organisms which have many more traits, leads to some of the misconceptions that we see, even in popular science writing and polemics. For example, in the debates over atmospheric CO_2 and its relation to climate change, some people in addressing the impacts have identified CO_2 as a fertilizer for plants. They rarely recognize that plants’ carbon fixation is done by organisms that also have traits related to water use, heat tolerance, resource allocation, etc.

Many scientific advances come when an unusual, extreme, or very simple system is studied, in which many confounding factors are fortunately not present. While this is often very productive in elucidating the nature of a phenomenon, or causal relationships within a study system, the unusual, or unusually simple, cannot automatically be taken as the “type” for all other systems. So, investigations that are based on such tractable and enlightening subjects beg all the questions of generalizability that any experiment in biology does.

There is another aspect of "population thinking" which is of importance for the growth of biological understanding, to wit, that populations are not composed of identical units, but rather have structure. In sexually reproducing organisms, there are members of different sexes in some ratio that may be relatively stable — or not. Not all members of the population are reproductive at the same time — very often because there is a diversity of life-stages or ages (or both) present, but also possibly for other reasons as well. Mating may not be random. Variation within the population for various traits may lead to differential survival or reproductive success. All these details are essential to the dynamics of a population not simply seen as a demographic entity, but as an ecological and evolutionary entity.

<u>5. Variability</u>. Obviously, another consequence of this thinking is that the mechanisms of life are at work in individuals that are unique, and the degree varies in importance depending on the kind of process being considered. Variation can be seen to be an important concept in learning about evolution; it is a commonplace that variation is the raw material of evolution. Prior to this, on a descriptive level, a biology student should learn to recognize variation in at least a few different species, not always easy, as except for trees and domestic animals, the untrained eye tends to regard all members of other species as essentially identical.[4] After all, it is in these variations that a great deal of taxonomy is rooted: How much variation, in what characteristics, is within a species, and how much demarcates between species? In a school setting, it would be important to explore when such variation has noteworthy consequences, not only in reference to natural selection, but also in reference to subjects as diverse as medicine and population dynamics. A synonym for "variability" is almost "individuality."

Having completed this quick tour of the "distinctives," we can examine whether and how the literature takes account of them, either as the direct object of study, or as an indirect feature, for example in the nature of the systems used in the labs, or in understanding the implications of the study for students' growing sophistication about these features of living systems.

E. Codes pertaining to biological distinctives

Codes were defined as follows (Table 2):

Nature of Biological Distinctive	1. Evolution Evolutionary process words are explicitly included or stated (e.g., populational thinking, phenotypic or genotypic variability, adaptation as a "state of becoming," form and function related to adaptation, data on different organisms or traits are explicitly compared, fitness). 2. Dual causation Phenomena are described as having both some environmental control and some

[4] "What!" snorted Bilbo, "You can't tell the difference between a hobbit and a man?" "To sheep, other sheep no doubt look different, " laughed the elf, " or to shepherds. But the study of mortals has not been our concern."

	genetic control (e.g., responses of different plant genotypes to environmental stress). 3. Complexity Phenomena are explicitly connected to the complexity of the biological system, phenomena are explicitly related across levels of organization (e.g., molecular triggers for organ behavior, individual ecology of a population of organisms addressed in an ecosystems context), may also include uncertainty/emergent phenomena. 4. Populational thinking Addresses some aspect of a population of organisms for its own sake, may include probability. Only select if not related explicitly to evolution. 5. Variability Addresses some aspect of the variability (of a population) of organisms, may include probability. Only select if not explicitly related to evolution.
Biological Distinctive is functional or historical	1. Functional Focuses explicitly on a biological phenomenon without any connection to the unifying ideas of evolutionary biology (may include adaptation, form and function discussed as current features of organisms). Focuses on "what" and "how" questions about phenomena (Mayr 2004). 2. Historical Takes into account some aspect of the unifying ideas of evolutionary biology (building explanatory narratives for evolution, drawing on various kinds of evidence – fossil, embryology, comparative anatomy, etc., adaptation, form and function). Focuses on "why" questions about phenomena (Mayr 2004).
Locus of biological distinctive	1. The author explicitly addresses the Biological Distinctives (BD) either in the theoretical framework or the discussion section of the article; the students do not. 2. The BD addressed by the lab experience is described as the context for the lab either in the article and/or in the students' experience. 3. The BD addressed by the lab experience is contained directly in the lab itself, that is, students address the BD by *doing* the lab activity. 4. The BD addressed by the lab experience is embedded in the work students do to analyze the *outcomes* of the lab.

Table 2.

F. Characteristics of the corpus related to the distinctives

1. Frequencies. For this research we considered evolutionary biology, dual causation, complexity, populational thinking and variability as five topics unique or distinct to the study of biology. These concepts are referred to as 'biological distinctives.' While others might be proposed, these, derived from the literature on the philosophy of biology, proved sufficient for the examination of this corpus. Of the 110 coded papers, 43 papers (39%) addressed a biological distinctive in their research (Fig. 24). Sixty-seven papers out of 110 (61%) did not address a biological distinctive. Out of the 43 papers that did address a biological distinctive, 29 papers (21% of the 110 coded papers) addressed only one category of biological distinctive. Nine papers (8% of the 110 coded papers) reported addressing two categories of biological distinctives. Three papers (3%

of the 110 coded papers) addressed three biological distinctives and two papers (2% of the 110 coded papers) addressed four biological distinctives.

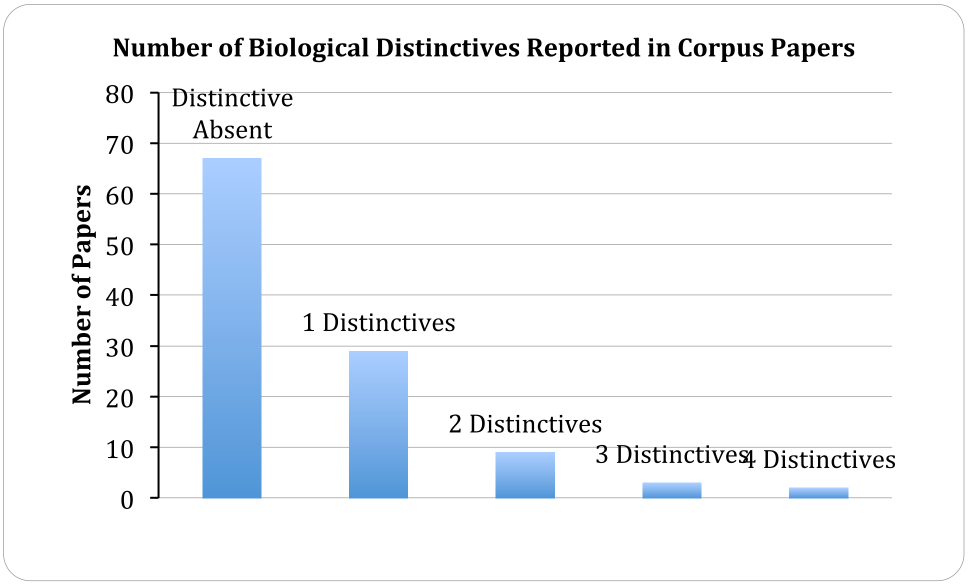


Figure 24.

By far the most commonly encountered of the "distinctives" was complexity (31 papers) (Fig. 25). This is not surprising, since a high proportion of the labs were designed to address ecological or genetic topics. One of the basic criteria for coding a study for complexity is the inclusion in the lab or the study of multiple levels of biological organization. Most of the genetics papers address the distinction between genotype and phenotype; the ecological papers either address trophic relations, or coupled human-nonhuman systems (for example in labs relating to conservation). The 13 papers which addressed populational thinking included primarily ecological and genetics topics, though at least two of the 16 evolution-themed papers also included a consideration of population processes. Four papers addressed variability, and one was coded as addressing dual causation. The way these themes were treated is discussed at more length below.

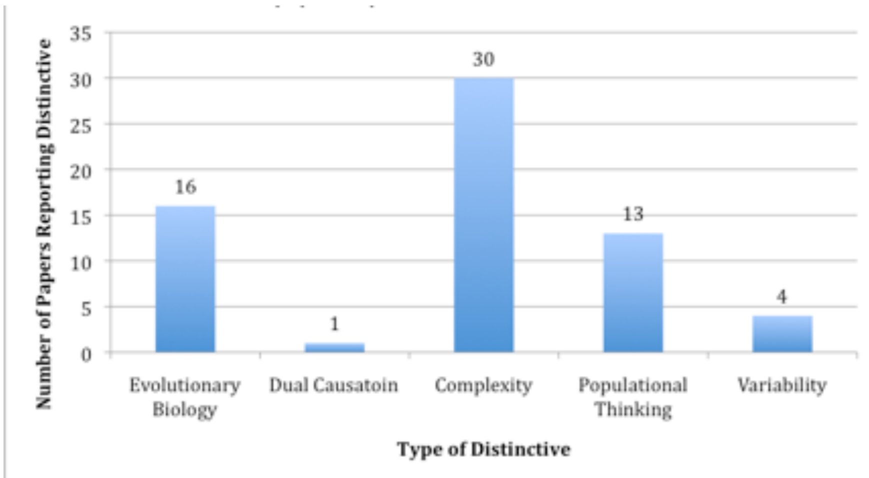


Figure 25.

We were also interested in whether a paper set in any kind of evolutionary context the phenomenon that was the focus of the lab. That is, even if evolution was not the principal topic of the lab, did the study reflect upon the evolutionary dimension of the topic? Out of the 43 papers that addressed a biological distinctive, sixty percent (26 papers) addressed the distinctive through a functional lens, that is to say that the biological distinctive did not make an explicit connection to evolutionary questions, even by way of cross-species comparisons (figure 26). Forty percent (17 papers) of the papers addressed the distinctive through a historical lens by explicitly taking evolutionary ideas into account. But the majority of these were focused on teaching evolutionary biology.

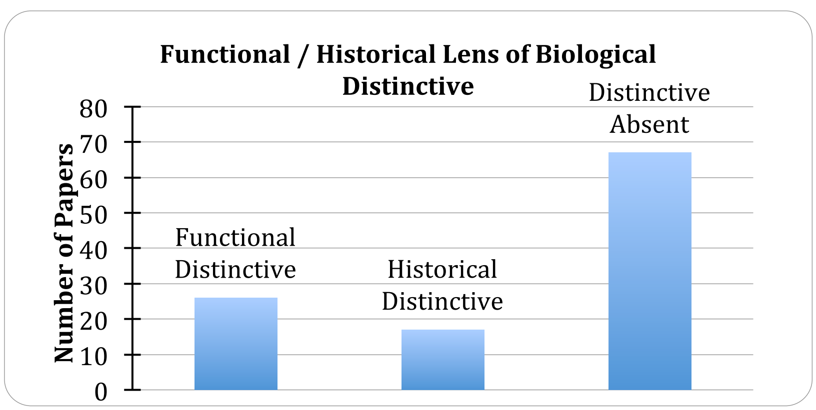


Figure 26.

In examining the corpus with respect to the "distinctives," we felt it was important to know where the distinctive was addressed, that is: Do the students engage the topic(s) as part of their laboratory, or is the topic only addressed by the researchers in their theoretical framework or discussion of implications for biology education? If the students engage with the distinctive in the course of the lab experience, then the study may provide information about ways to support students' learning to reason about, and investigate, phenomena that present one or more of these key features of biological systems.

Of the 43 papers coded for inclusion of one or more of the distinctives, twelve reported that the biological distinctive was addressed by the author(s) alone, and therefore in 31 the students also engaged with them (Fig. 27). In some cases (11), the distinctive was included as part of the context-setting or preparation for the lab. In 39 papers, the lab activity itself required the students to address or make use of the distinctive; and in 14 studies, the students were required to address the distinctives explicitly, as part of their sense-making about the lab activity.

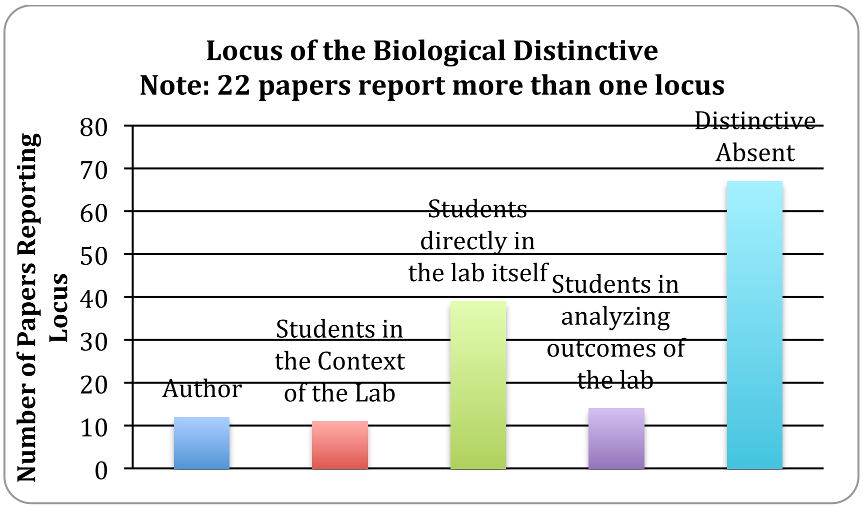


Figure 27.

2. Thematic analyses

a. Evolution

Sixteen documents were coded as addressing evolution. In all but 4 of these references, the aspects of evolution addressed are mentioned only briefly. For example, in Taraban et al., the content of the intervention includes "DNA structure and function, protein synthesis, and natural selection," in the context of biotechnology. When Grace and Ratcliffe analyze the values that students use to debate conservation scenarios, they note that evolution is not mentioned. Slack and Stewart focus on "transmission genetics" and generational thinking, and clearly set their work in the context of an evolutionary frame. Brisbin focuses narrowly on the construction of phylogenetic trees, but the reader cannot know how this is related for the students in the study to other aspects of evolutionary theory (for example, processes of speciation).

In a few cases, however, the authors are more explicit about specific elements of evolutionary theory involved in the laboratory activities they are analyzing. Beardsley () studies middle-school students' capacity to learn to include 7 principles: [a] there is competition within and between species, [b] not all offspring survive to reproduce; [c]

survival of offspring is limited by environmental factors; [d] individual variation is related to individual survival; [e] there is specialization of species in particular niches, and that environmental change can affect these niches; [f] of the variation seen within species, some is genetic (arising from mutation or recombination); [g] selection has non-random effects on differential survival. Geraedts and Boesma (2006) and Passmore and Stewart (2002) are equally explicit in the version of evolution by natural selection that is being taught to the students in their study. Finally, Gallucci's literature review (Gallucci 2007) provides a broad overview of the current core theory of evolution; while the study examines a range of possible conceptual changes in the subjects, the learning of evolution receives some careful consideration, as Gallucci finds that when students understand variation in populations, this enables a "conceptual cascade" of understanding about other aspects of evolution, especially natural selection. Variation is thus a "keystone" concept, in this case.

What can we learn from these papers about the role of labs in facilitating understanding of evolution, or supporting evolutionary thinking? First, it must be noted that these pieces were "credited" with addressing evolution because the topic is mentioned, not necessarily because the laboratory being studied addressed evolution explicitly. For example, Marbach-Ad and Classen (2001) are intent primarily upon an approach to improve college students' questions. They provided a stimulating context within which to encourage questions of many kinds, and also to help the students learn how to move from question to investigation in the inquiry-oriented labs the course employs. Their taxonomy of questions includes a category requiring an evolutionary or functional answer.

In another case (Slack and Stewart 1990), evolutionary topics arise in the context of lessons aimed at the development of genetic problem solving. In this study, students are presented with problem spaces that include computing and explaining genotype and phenotype changes from one generation to another. Here, however, the focus was on genetics, rather than processes of evolution, and so ecological or population-genetic dimensions were not explicitly addressed.

Similarly, Hoesen and Nowiki (2001) address evolution, but only in the context of a broader innovation in their biology class — having their students apply each topic in the biology course to an organism they chose at the beginning of the semester.

Three studies target evolution directly: Passmore and Stewart (2002), Geraedts and Boersma (2006), and Beardsley 2004. Geraedts and Boersma engage the students in a telescoped re-enactment of history, in which the class successively learns about and deploys Lamark's, and Darwin's theories as explanatory tools with which to make sense of problems posed by the teacher. Starting with a case scenario in which acquired characteristics might possibly be inherited, students are led through a series of questions that are designed to bring to bear their understanding of basic genetics and other biological topics, and guide them to a reinvention of a version of neo-Darwinian theory. Finally, they use a simulation game (designed by Stebbins and Allen) to learn the basics of natural selection. This approach, while highly structured, nevertheless requires students to reason on the basis of theory, and to confront limitations of their explanations revealed by phenomena in the cases or simulations presented. The authors suggest, on the basis of written post-test answers, written answers produced during some of the stages of the intervention, and pre- and post-intervention interviews, that the process helped students increase their incorporation of modern evolutionary thinking, even though there remained some lacunae in the understanding even of the students classified as "neoDarwinian." Since there was, however, no written pre-test, nor other basis for a quantitative comparison pre- and post-intervention, we cannot tell from this study how significant an effect was produced by this approach.

Beardsley (2004) is one of the few papers in our corpus to address evolution in middle school. The gist of his paper is foreshadowed by his title: "Middle school students learning evolution: Are current standards achievable?" He describes a study in which middle school students engaged in a series of activities — discussion activities, hands-on investigations of variability in populations, and some direct teaching — designed to address the State of Washington's "Essential Academic Learning Requirements" for evolution in 8th grade. While there was a positive trend in students' understanding of core ideas about evolution and natural selection, "only a mean of 25% met the standard," a percentage similar to that seen in some other states. Beardsley notes that these results come even with the use of an intensive unit, implemented by highly knowledgeable teachers over a period of time probably not typical for a middle school's curriculum on evolution.

Passmore and Stewart provide an interesting contrast to Beardsley. Their study is intended to be a "power test," with a carefully constructed 9-week course. Theoretically and scientifically deep, the course has students test the explanatory value of Paley's "divine design" theory, Lamarck's view of evolution, and Darwin's model of natural selection. The authors find that by the end of this course, students are able to develop quite rich analyses and explanations of evolutionary phenomena presented to them. The authors acknowledge that the course as is could not be widely implemented, yet they provide strong evidence that students who have a range of academic achievement can understand and employ the key elements of modern evolutionary thinking.

What, then, are some lessons to be learned from these papers?

1. There is a paucity of careful studies of laboratory experiences that help students understand evolution. There are a few studies in which students gain an understanding of some of the elements, such as natural selection, or the process of inheritance of traits and some of the basics of genetic variation. The "craft literature" and the curricular literature are full of examples of activities which instructors or writers have devised and tested in their own settings. The research literature essentially has not explored the value of these innovations.

2. The studies that we have seen are small-scale, intensive studies, whose generalizability is not clear. Passmore and Stewart (2002) and Beardsley (2004) are explicit in their judgment that, though their interventions lead to students' gaining significant understanding about evolution, they were not likely to be replicable in most classrooms (and perhaps not by most teachers). Given this, Beardsley's question about the feasibility of achieving the prescribed content standards on evolution can be asked about standards across the country, and across grade levels. There are useful studies evaluating the content standards for evolution, for example state-by-state, but evidence suggests that having good standards in this area does not mean that the content is being well-taught, nor learned.

3. The studies in our corpus that address evolution directly have taken a broad approach, seeking to involve the students in understanding evolution as an explanatory theory with several basic assumptions, and to help them gain some sense of its increased adequacy as an account of natural phenomena, in comparison with prior theories. The "lab experiences" were appropriately complex, and involved multiple elements, such as laboratory activities, discussions, simulations or thought experiments, and reading or other exposition. If one agrees with Passmore and Stewart that "scientific practice is discipline specific...curriculum should therefore take into account the ways in which scientists operate in their fields." (p. 187), then an adequate understanding of evolution must be buttressed by an acquaintance with a very broad range of ideas, methods, and phenomena in biology. To that extent, an "integrated lesson" approach (NRC 2006] seems warranted.

4. Understanding evolution, and applying it to phenomena, requires both considerable acquaintance with biological facts, and some sophisticated reasoning, including the use of comparative information, and several kinds of inference. One theme that is addressed in the "evolution" papers in our corpus is the importance of reasoning ability. One paper, Johnson and Lawson 1998, directly tests the relative importance of reasoning ability vs. prior knowledge as factors of achievement in an introductory college biology course, and finds that across the course (not only in relation to evolution) "reasoning ability but not prior knowledge or number of previous biology courses accounted for a significant amount of variance in final examination scores." (pg 89). Other papers in the corpus address the reasoning challenge as well. Slack and Stewart (1990) analyzed the problem-solving strategies of 30 high school students solving problems posed by a computer program which requires students to "plan experiments, generate and interpret data, and reason from effects (phenotypic data) to causes "genotypic data." (p. 55). Their careful study showed that students had a hard time developing explanations both warranted by data and constrained by theoretical considerations. Moreover, the authors identified "genetics specifics ways of thinking" which students lacked: "genotypic thinking, generational thinking, and ability to distinguish between an inheritance pattern and a modifier." (p. 64).

Obviously, genetic reasoning is only one piece of the evolution puzzle. Another approach to the challenge of biology reasoning is represented by Lavoie (1999), which studied an intervention with high-school students in a biology course explicitly structured around the "learning cycle" of Karplus, which comprises "exploration, term introduction, and concept application," with use of hands-on activities in at least the first phase. The additional intervention was to add an explicit step of generating hypotheses before the "exploration" phase, and then using the hypotheses to focus subsequent work. This step resulted in significant gains in students' learning of science skills and concepts, and in their engagement with the course.

5. Interaction is important. Four of the papers make extensive use of technology of some kind, either to generate data for students to reason about, or to simulate systems that students could explore interactively, or both. Other studies used physical simulations (e.g. the Stebbins "selection game"), or extensive case materials that provided resources that were rich enough to support students' follow-on questions and investigations. In all cases where success was reported, whether in conceptual understanding or reasoning ability, the lab activities involved interaction — between students and between students and

simulations; and several also enabled reasoning to be checked against additional information. Not surprisingly, given the challenges of observing natural selection in a population of organisms in lab or field, the use of living systems was highly constrained and primarily illustrative (e.g. to the measurement or other observation of variability in a population), and embedded in a larger narrative framework. Yet it is worth noting that biology educators have devised methods for studying populational features, including selection and evolution, for example in microbial microcosms (Brockhurst 2010).

6. Technology helps, but time may be the secret ingredient. Several of the papers in our corpus made use of simulations to generate data (esp. genetic data) for the purposes of teaching evolution or at least multi-generational genetic transmission (phenotype/genotype data). This accords with a national trend away from hard-to-mount "wet" labs, and towards virtual labs in all sciences. These tools, when well designed, have all the virtues of virtual labs (low cost, ease of set-up, ease of replication, possibilities for innovative and informative student representations of their knowledge), with their drawbacks as well. Especially given the complex nature of the ecological and population-genetic aspects of evolution, it seems possible that such systems can introduce misconceptions which may be difficult to identify and correct, so that whether one uses paper narratives, Web quests, physical simulations, or computer environments, evolution cannot be taught effectively without considerable time being spent in the exploration of multiple cases, with enough time spent in developing explanations and arguments that the teacher and the students can examine and improve their biological reasoning, based on an adequately rich set of phenomena, and both observational and experimental investigations, at least within very rich simulation environments, but ideally with a diversity of resources, including living systems.

However, while all these considerations are suggested by the literature reviewed, they can stand as conjectures only, at this point, given the small number and small size of the studies.

b. Populational thinking

In our corpus, 13 papers were coded as dealing with populational thinking. Not surprisingly, several of these overlapped with others in the collection. What meanings does "population" take in these papers?

Perhaps the clearest way to put it is, that populations are seen in these papers, universally, as demographic collectives, rathern than evolutionary/ecological assemblages. That is, populations are represented as numbers of individual units, with no reference to genetic, age, stage, or even sexual structure within the population. Typically, population numbers serve as dependent variables in a system in which other variables are being manipulated.

In 7 of these papers, "populations" are studied by means of computer simulations of some form. In one paper, students observe their populations in the field, but primarily in qualitative terms.

In a few cases, such as Faryniarz and Lockwood (1992), the conceptual setting is an ecological one — in this case, students were given "many parameters associated with populations and communities…kinds of growth, carrying capacity, migration, type of habitat, minimum breeding density, escape rates, hunting pressure, and predator/prey ratio." (pg. 460). The goal is to manage a deer population — presumably keep it within some acceptable range. Manzanal et al. (1999) have students analyze various facets of a microcomuter-simulated ecosystem, including the effects within a trophic chain of changes in numbers of one species occasioned by changes in the number of another species in the chain.

In other papers (at least 6 of the 13), populations are mentioned, but either used simply as outcome variables or not at all in results and analysis. In general, therefore, "population" in these papers represents "the group of organisms used for this study."

One paper, Geraedts and Boersma 2006, reports on a novel course designed to teach evolution by natural selection, in which Lamarkian or creationist frameworks are challenged by data in a structured format which guides students to the derivation or re-discovery of Darwin's theory as the best fit with the facts. In their discussion of the conceptual background, the authors give quite a clear evolutionary account of populations, and the course materials do engage students with populations (in highly simplified form) as evolutionary/ecological units.

Finally, an additional paper, Beardsley 2004, middle school students do an activity on variation within populations, in which they measure leaf-length of ten mature leaves on one of three different plant species found near the school (p. 607). Although students were asked to arrange these data in histograms, and make various predictions about the present and future populations, the biology underlying any such exercise is not clearly articulated, and it may well be that the activity is one of many that conveys the notion of variation among individuals, without any way to understand the meaning of it ecologically or evolutionarily.

In sum, the papers in which the idea of population thinking occurs give no evidence that the laboratory material, nor the conceptual surround (set-up or sense-making, curricular setting) serve in any way to help students understand populations as ecological or evolutionary entities, with genetic or other kinds of structures. Furthermore, none of the papers provide evidence that students do any work with the characteristics of actual or simulated populations, except as numbers of individuals.

c. Dual Causation

Only one paper was coded as addressing this "distinctive," the paper by Slack and Stewart (1990) on students' problem-solving strategies in doing "realistic genetics problems." The reason for the coding was the authors' inclusion of problems that would require students to take into consideration "modifiers" of simple Mendelian inheritance such as pleiotropy, lethality, or sex-linkage. The paper, which represents a careful examination of what kinds of causality students in high school will tend to use, strongly suggests that laboratory activities relating to multiple causality will be highly challenging for students to engage with (and for educators to design). The experiment made use of a genetics simulation program which provided a rich source of data for the students to explore, allowing them to repeatedly form crosses on successive generations of "field collected organisms." The results of this paper suggest that in helping students in a general, introductory course understand multiple causes for biological phenomena, a cyber-enhanced environment may be almost indispensible in offering the ability to replicate "experiments" enough variables to uncover multiple contributing causes, within the time constraints of the typical curriculum.

A search through other papers in the corpus, whose topics included genetics and the genotype/phenotype distinction found no examples in which the labs being studied engaged the students with the complexities that contribute to the development of phenotypes.

d. Variability

Four papers in our corpus were coded as involving the concept of variability. In the first of these, Grace and Ratcliffe (2002), the concept comes up only to disappear again. The authors report on an activity in which students identify the values which they believe to be important in decisions about species conservation. For the sake of comparison, the authors asked a number of conservation biologists to name the values they think important, and two among the many mentioned were 'variation between species' and 'variation within species.' Teachers in the study did not identify these ideas as likely to be invoked by students, and indeed the students did not name them.

Geraedts and Boersma (2006) describe a sequence of structured discussions designed to help students "reinvent natural selection." The questions, with some simple props (e.g. diagrams, images), confront the student with their own ideas about the processes of evolution, and with evidence that bears upon these ideas. One of the earliest of the questions addressed takes up variability within in a population. According to the authors, 90% of students were certain that there must be individual variation — probably a lot — within populations of any organism. The activity continues to build up layers of reasoning enabling the students to compare the adequacy of Lamarkian vs. Darwinian conceptions of evolution by natural selection, and the end result is that the students come to see for themselves why Darwin's view was superior in its ability to deal with the facts of natural history. The curriculum, therefore, does not by its nature provide space for students to explore actual variations in groups of organisms.

e. Complexity

Thirty papers are coded as addressing "complexity" in our corpus, making this the biological distinctive most frequently present in papers. There were relatively few subject domains addressed in these papers. In 14, the subject matter was ecology or environmental science; in 8, the subject was genetics. Two papers addressed biotechnology, three some aspect of cell physiology, one the circulatory system, and one placed the whole organism at the center of a year-long "applications" strand for a biology curriculum.

"Most areas of study that are at all worth our attention entail far more complexity than is acknowledged in our curriculum; and, further, people's intellectual engagement, when they are given the chance to pursue these complexities according to their own lights, is extraordinary. Our challenge as curriculum developers is to find the ways to engage learners, young or old, in the complexities of the areas we think it is important for them to know about" (Duckworth, 1991, p. 23).

Recall that we coded for "complexity" with the following rubric: "Phenomena are explicitly connected to the complexity of the biological system, are explicitly related across levels of organization, may also include uncertainty/emergent phenomena" (Table 2). Two topics stand out as subjects for these lab activities, genetics (11 studies) and ecology (16). Cellular and molecular biology, organ systems, and evolution are all represented by far fewer studies (3, 2, and 3 respectively). In the largest number of cases, the "complexity" present was coded because the lab addressed more than one level of biological organization. In studies with an ecological focus, the phenomena studied most often involved trophic relations, e.g. food webs or food chains. In genetics studies, the complexity arose from the genotype-phenotype relationship, and the "abstract" nature of the processes at work.

Few of the 31 papers coded for "complexity" addressed that topic explicitly; rather, they assume or assert that the topic of the study is a complex one, and they may address this complexity in their rationale ("genetic systems are complex...") without actually characterizing the nature of the complexity that they are responding to[5]. Complexity is taken to be inherently a source of difficulty for the learner.

However, the pedagogy of some of the labs being studied is designed to reduce complexity, or to scaffold the students' encounter with it; complexity is both a feature of biological systems, and a problem for the learner. In some cases, the problem space (with or without a computer-based representation) is designed to ensure that all parts of the system being studied (for example, in the respiratory system) are kept in mind, so that the student can less easily develop a partial conceptualization of the system under study. For example, Law and Lee (2002) describe an interactive genetic microworld system that is designed to scaffold students' understanding of simple Mendelian genetics. The students can run breeding experiments with cybernetic rabbits, and the system provides them with state information (e.g. genotypes being crossed), as well as ways for them to manipulate inputs and even some of the genetic "rules" to produce outcomes needing analysis.

Such lab settings, especially if computer-mediated, basically provide deterministic systems to manipulate. For example, Faryniarz and Lockwood (1992) used a simulation program to help community college understand conservation or resource-management challenges (e.g. the management of deer populations). It is evident that it is not realistic to experiment with deer population regulation in the classroom, but the management element in the challenge renders some aspects of it both realistic and engaging. The constrained nature of the model directs students' attention to key elements in the system; but the rule-based nature of the system limits the realism and the complexity of the representation of the human-habitat-deer system. As with the genetics systems, the deterministic and constrained nature of the systems has notable advantages: the ability to conduct "experiments" on otherwise inaccessible systems, and to conduct replicate studies frequently and quickly. This "repeat" function, especially if the system is interactive, can in itself facilitate the development of students' mental models, providing them with patterns from which to reason.

---

5 However, Yoon (2007) does offer this: "complex systems exist: when any given number of interconnected elements, parts, or individuals communicate in non-linear ways. The pattern of interactions forms a collective network of relationships that exhibit emergent properties not observable at subsystem levels."

It is interesting to note, however, that the aim of these studies can best be characterized as conceptual understanding, that is, the mastery of particular concepts. Hence, the deterministic approach makes sense, as it places careful constraints on the ideas that students can construct, and in some cases is designed as well to eliminate specific identified misconceptions. These studies deal with complexity by minimizing it for pedagogical purposes

A contrasting approach engages the student with the complexity, even if it is mediated or moderated by instructional materials or methods. Buckley (2000) defines a complex system as one that has multiple parts, with multiple interactions, operating at multiple physical and temporal scales, and composed of interacting subsystems. This is her characterization of the complexity of the human circulatory system, and on her analysis this complexity is a key source of difficulty for students learning the material. This conclusion follows on from her definition of learning, which involves students' building mental models; and she deployed a multi-media system to support this development, as students developed a mental model of the structure and function of the heart. The system provided multiple representations of the heart's structure, function, behavior, and relationships in the system, and the two students Buckley followed in detail developed increasingly rich and accurate mental models (to use her terminology) of the heart. Buckley judges that this knowledge would facilitate an increasingly rich and adequate mental model of the whole circulatory system, if the students continued developing deep understandings of the sub-systems. This reductionist approach is in the service of a holistic or systems strategy, and the multiplicity of representations, facilitating students' learning at their own pace, is central to it.

An alternative approach represented in the literature in our study involves the use of modeling and the building of systemic representations by the students themselves to explore the dynamics of complex systems. Many of these are computer-aided, but also have some "real world" components, or are entirely non-digital. When embedded in an understanding of scientific reasoning and experimentation as theory-building and testing, the use of multiple tools can add power to the learning system.

Verhoeff et al. (2008) have such a view of science, but they also are concerned to represent biology from a systems point of view, in which macroscopic (organismal) phenomena are linked explicitly with phenomena at lower scales of organization (including at the cellular and organ-system level). They were seeking a pedagogical approach that draws students to look at one "thing" — digestion — in this systematic way. Their lessons move through several cumulative stages of drawing, observation, creation of 3-D physical models of cell and organ structure, they then used a computer system to build diagrams of the digestive process from cell to organism (or *vice versa)*. The authors assert the work from observation to representations (an iterative process) provided the students with a concrete understanding of the components of the system at each level (e.g. the organelles in plant and animal cells, and their spatial situation within the cell) which was necessary for a systems understanding of the phenomena. It also appeared that the thoroughgoing use of the systems approach resulted in students' inclination and capacity to apply it to new systems (e.g. breast-feeding).

Stratford et al. (1998) also engage students in building models, this time of a stream-system, using a simulation environment which provides a range of variables from which students can choose in constructing a simulation to explore a question (e.g. the effects of eutrophication, weather effects upon water quantity and quality, and so forth). The system does not specify the relationships between variables, and allows students to set values, and then "run" the simulation, providing feedback (data) about the state of input and output variables. In the class analyzed in their paper, the students had participated in a prior stream study near their school, learning to map and measure the water course, and take data on various measures of water quality. The simulation study did not have the students testing their models against the real world, but the students' previous experience with an actual stream system provided them with some understanding (whether systematic or pre-theoretic) of such systems. Therefore, the simulation seemed realistic or as the authors say, authentic, and their real experience was a help in understanding and using the modeling system. Yet insofar as the paper portrays the students' reasoning (and this is an important focus of the study), it appears that the field experience's impact was on the students' understanding of key variables, such as dissolved oxygen. In none of the reported discourse do students compare their model to the stream they have come to know during the preceding 3 months. Even the most sophisticated of the student teams seem to be doing their reasoning within the micro-world of the simulation. It may well be that the modeling helps understand key systems behaviors, but it remains to be seen whether it helps students analyze actual systems.

Finally, Magntorn and Helldén (2007) present data that suggest that detailed "natural history" knowledge of a system facilitates the application of systemic understandings in new settings. This in some sense fits with the Verhoeff at al. paper described above. Magntorn and Helldén introduced students to basic ideas of ecosystems, including energy flow and cycling of matter. The presentation of these ideas moves past an abstract and schematic engagement, commonly seen in textbooks. Instead, the students learn about the species of the system, so that trophic relationships are seen concretely embodied. They are taken to a field site, to learn to see these organisms, and watch them interacting, so that they see evidences of the systems processes at work. The students worked with the ideas in the lesson, and their own reasoning, using a variety of representations, from textbooks to concept maps to the building of microcosms.

Not only did the students learn about particular species which they did not know before, but in the process they seem (according to the authors' data) to have understood the ecosystem processes embodied by these organisms to the point that, when taken to a different system, the students were able to "read" the new environments. Concepts included photosynthesis and primary production, succession, decomposition, causal relationships between abiotic and biotic factors, nutrient cycling, and others.

Another way to describe the learning in this study is to say that the students have developed a theoretical understanding about habitat processes, functional groups, and species characteristics that they can then bring to bear when confronted with unfamiliar systems. They were prepared to look for tropic relations, for guilds of organisms (e.g. detritivores), for the partitioning of resources, etc. The concreteness of the work they did in their first system provided the foundation for this theoretical understanding (however naive). As the authors write, "A critical feature constraining the transfer of reading nature seems to be the lack of species knowledge. By the pond the students seem to know what to look for but find it difficult to transfer their knowledge of structures since they do not recognise (sic) the organisms by names as they do in the forest" (Tatar and Robinson, 2003).

Thus, some of the 31 studies in our corpus coded for "complexity" seek to address the complexity of the systems they use by constraining it — limiting the number of variables examined, or reducing the system to a very simple model. Others provide various kinds of scaffolding to build up students' acquaintance with key, "memorable" features of the study system. Some, however, such as the last 4 discussed, address the complexity without seeking to eliminate it.

Many of the studies incorporated elements that enabled the students to become more familiar with the elements of a complex system, though only some authors explicitly note the importance of gaining detailed knowledge about the components of the system (e.g. organisms, environmental features) under study. Some did this through models, which allowed students to see repeated events in which inputs and outputs could be manipulated, and in some cases to construct elements of the model (Slack and Stewart 1990, Hogan 2007, Tsui 2004, Law and Lee 2004). Others used field activities (Prokop 2007, Magntorn and Halldén 2007, Hamilton 2007, Manzanal 1999). Still others used a resource such as video disk libraries (Leonard 1992), or materials-supported interactive activities with a strong component of reasoning and group discourse (Seethaler 2004, Jimenez 2002, Geraedts 2006, Yoon 2007). In all these cases, an important feature is the ability of the students to articulate provisional understandings, check them against some external authority (a data set, data generated by a simulation, visual or other references, peers), and revise their understanding.

The modeling approach, whether students build models in some external representational medium, or merely construct mental models, expressed primarily in words, seems to provide a powerful approach which preserves some level of complexity in the study system, by providing an iterative process of theory building and interaction with the living system, or a simulation of it, and emergent properties can be addressed in this way as well, though they are little discussed in the studies at hand. In these cases, however, "signal" is still ensured, and the study system does not require of students the challenge of understanding how to separate signal from noise — that is, the effects of natural variability, which would require larger data sets, mathematical tools for analysis, and a theory which would provide some coherent account of the reasons for variation.

One other strategy seems noteworthy, which is that seen in Hoese 2001). Hoese here reports not on a specific laboratory exercise, but a yearlong strategy in which the students are to relate every aspect of the course throughout the year to a specific organism. This simple strategy has the advantage of requiring the students to take whatever phenomenon they are learning about, in any level of biological organization, and relate it explicitly to the organism, which is the unit of ecological and evolutionary events, and in some sense the functional focus of cellular and physiological processes. This then provides a lens through which to impose order on the complexities of the subject matter, and with which to interrogate their developing mental models of various aspects of biological science. This then provides the systems focus advocated by most of the authors in this corpus who explicitly address cognitive strategies for dealing with complexity — and it is notable that Hoese himself does not name this as a target of his strategy.

G. Conclusions regarding biological distinctives

This part of our literature synthesis was undertaken to address the question, "To what extent do the labs represented in the research literature engage students with the biological distinctives?" The short answer is "To a limited extent— we found evidence of such engagement in just 43 out of 110 studies." But how shall we interpret this finding?

In the first place, it is well established that a significant proportion of lab experiences are designed to illustrate a specific concept (e.g. diffusion through a semi-permeable membrane, territorial behavior in male sticklebacks, how to do a mark-and-recapture estimate of animal populations, how to use a microscope) (NRC 2006). Another proportion of labs are intended to illustrate a concept (Mendelian inheritance) or enable students to apply something they've learned in a practical setting. In these cases, the intent is to ensure that students get a clear understanding of some specific facts or scientific results — propositional knowledge, sometimes also to include some of the implications of that knowledge. If the lab is well coordinated with the "text" curriculum, the text is relied upon to supply the meaning, context, and implications for the laboratory activity. So while a lab on some specific feature of physiology could motivate the consideration of systemic implications of the phenomenon under study, labs often don't.

The studies in our corpus reflect this general picture. These studies were conducted during the "standards era," and the need to reliably address science standards in preparation for high-stakes tests certainly creates pressure to design labs that will lead to predictable outcomes. Yet such labs have always been part of the curriculum, even in eras when inquiry has been highly valued (DeBoer 1991), as they serve valuable purposes. And there are a few examples in which students are confronted with more complex and phenomena to study, as described above. Nevertheless, we had hypothesized that most studies would not address the scaffolding of biological reasoning to involve students with biological systems characterized by complexity, variability, populational and evolutionary elements. This hypothesis is borne out, as almost all the papers in the corpus, even those coded for the presence of biological distinctives, significantly simplify and constrain the systems that students are to study. Moreover, there is no indication that the constraints and simplifications are more characteristic of labs for younger students than older ones. Thus, the lab designs do not represent "scaffolding" in the sense typically used in educational research, since the "scaffolds" do not fade away (insofar as we can tell from the research papers), challenging and enabling the student to inquire into biological systems with all their diversity and complexity.

10. STUDENT OUTCOMES

Codes pertaining to student outcomes were defined as follows:

1. Science subject matter content knowledge
2. Scientific reasoning/making sense
3. Knowledge of the process of science (e.g., designing an experiment, as well as science process skills such as dissecting a frog, making careful observations)

4. Knowledge of the nature of science (defined by the AAAS in the Atlas (2001) as follows: "The nature of science involves the basic values and beliefs that make up the scientific world view, how scientists go about their work, and the general culture of the scientific enterprise.")
5. Motivation and Engagement – Describes an affective experience such as engagement, enthusiasm
6. Attitudes about science, about the scientific enterprise
7. Participation in class
8. None
9. Other
10. Student perception of what learning is like in the environment of the science classroom

Table 3.

Not surprisingly, given the percentage of articles that stated learning content knowledge as one of the instructional purposes of the lab, 95 papers (86%) reported content knowledge outcomes (see counts on next page, Figure 28). Almost half (48%) reported on students' scientific reasoning and almost a third (31%) reported on science process skills. Exactly half (55 papers) reported some measure of student affect (motivation and engagement, "participation" in class, or perception of the science classroom environment), while 23 (21%) measured student attitudes toward science. Interestingly, only 8 (7%) explicitly reported on knowledge of the nature of science. Note that totals are higher than 110 because most authors reported on more than one student outcome. Student outcomes are discussed in more detail below.

A. Content knowledge and scientific reasoning

Authors researching the impact of novel lab applications generally want to know about the labs' efficacy as well, and therefore seek to establish whether or not students have learned from the lab being studied, almost all measured by means of some form of pre-post assessment. For this reason, the high percentage of papers reporting on student learning outcomes is not surprising. More interesting to us is the subset of papers in which researchers not only measured the impact of labs on content knowledge but also on reasoning. Thirty two percent of authors (35 papers) reported both outcomes.

In this section, we briefly describe i) the kinds and duration of activities students engaged in to produce reasoning outcomes, and ii) what kinds of reasoning outcomes were described as a result of these activities.

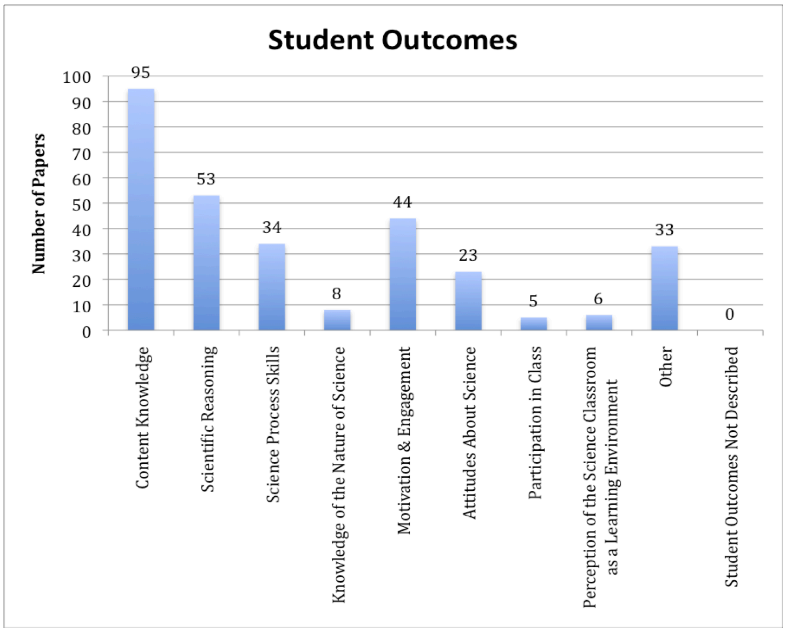


Figure 28.

i) Kind and duration. Students were engaged in extended lab activities in just under two thirds of these studies (21 of 35 papers), defined arbitrarily as a duration of more than 5 lessons. Besides being given the opportunity to spend at least a week engaged in the learning activity, topics were presented in a generative context in almost all of the 35 papers, either involving modeling, or explicitly scripted as "problem-based" or "inquiry-based" activities, or focused on living systems:

Context	No. of studies
Modeling	14
Problem-based	5
Inquiry	11
Living systems (Ecosystem or Microcosm)	5

Table 4.

Thirteen of the fourteen modeling studies involved software; one involved the construction of physical models. The most common domains studied through modeling were either genetics (e.g., focusing on genotype-phenotype interactions or gene expression) or systems (e.g., circulatory, ecological). The problem-based studies either focused on environmental planning (3 studies) or on genetically modified organisms (2 studies). The goal for seven of the 11 "inquiry," or "constructivism" papers was to contrast how students learned in inquiry environments with those learning in "traditional" settings.

What do we mean by the description of context as "generative?" These topics and means of studying them are all generative in the sense that they placed cognitive demands on students that went beyond the arguably "simple" task of learning science content, by means of a lab design that engaged students in iterative design-as in the modeling activities-or the process of science-as in the problem-based and inquiry-based studies. So, for example, in studies that included extended lab activities, students were required to manipulate and observe the effect of three independent variables on growth rates in a simulation (Huppert et al., 2002), critically analyze sources of information and authority in a conservation decision (Jimenez-Aleixandre and Pereiro-Munoz, 2002); or move "to an increasingly interpretive role using experience and some information learnt at school to explain that which they observe[d]" (Tomkins and Tunnicliffe, 2001, p. 803) when conducting extensive observation of brine shrimp ecosystems. Even in brief lab activities, students were required to engage conceptually, for example, by sharing internal mental models verbally and through visual representations (Para and Sarapuu, 2006) or to reason about inputs and outputs as they constructed quantitative ecological models using STELLA software (Hogan and Thomas, 2001).

ii) Reasoning outcomes. What types of reasoning were reported as student outcomes in this subset of the corpus? The categories listed in the table below (Table 5)emerged from the studies themselves upon closer examination of the "Student reasoning" results. Note that the total is larger than 35 since many authors described more than one reasoning outcome:

Type of reasoning outcome	No. of studies
Justifying claims/formulating and defending explanations	11
Critical analysis	9
Generalizing	6
Modeling own assumptions	4
Reasoning about relationships	4
Generating hypotheses	3
Revising models	2
Reported as "reasoning" without further detail	12

Table 5.

Twelve authors merely reported that students showed gains in reasoning without providing enough detail for us to be able to categorize the type. While these authors provided detail about how the lab activity of interest was structured to support student reasoning, it was frequently difficult to glean information on how students were actually reasoning as outcomes data. That being said, the most common reasoning outcome for which details were provided was that of justifying claims to classmates, either by pointing to scientific evidence or by providing warrants for claims. Also included in this category was discussion with classmates to provide and defend "explanations." So, for example, Geraedts and Boersma (2006) describe how students argue their interpretations of a proposed mechanism for how species change by reasoning from prior knowledge against a Lamarckian mechanism: "the inhabitants [of Australia] have remained white. They should have been dark, as they become darker, because their parents were tanned by the sun" (p. 853). Windschitl and Andre (1998) investigated the effects of a "constructivist" learning environment on conceptual change using a computer simulation of the human cardiovascular system as instructional tool. They suggest that, because students could create and test hypotheses, the simulation was

"more effective in changing alternative conceptions presumably because it highlights [to the student] the path of reasoning used to arrive at a conclusion" (p. 157).

Critical analysis, identifying key issues in a case, key concepts in an activity or key components of a system, was the next most frequent reasoning outcome described. Three examples that give a sense of the range of outcomes included in this category were identifying key issues related to applications of gene technology (Lewis and Leach, 2006), extended student opportunities to "read nature" in pond and forest ecosystems, that led students to relate concrete organisms to abstract functional groups and further to abstract ecosystem processes (Magntorn and Hellden 2007), and analyzing an ecosystem flux model to understand its dynamics (Hogan and Thomas, 2001).

Some papers included other measures related to student cognition and thinking. These papers included measures on: the ability level of students; student cognitive structures; misconceptions; and the transfer and retention of knowledge. Ajewole (1991) looked at student ability level in relation to student attitudes. He found that the "high-, average-, and low-ability" students in the discovery oriented biology class had a more favorable attitude towards biology than the students in the control class (p. 401). In studying misconceptions, Law and Lee (2004) found that "the science stream students [as opposed to the humanities stream students] had a tendency to interpret situations and observations on the basis of not only their own intuitive ideas, but also the concepts learnt through the science curriculum, resulting in a more diverse collection of misconceptions" (p. 121). Zohar (1996) found that "students who achieved 100% valid inferences in the late interview of their first problem were able to transfer the strategies they had acquired to a new task… The results from the retention problem indicate that those nine students retained their acquired strategies across time" (p. 6).

B. Motivation and Engagement

The measures reported by authors for motivation and engagement varied widely. Measures included motivation, engagement and other affective responses such as student ability to stay on task, what students liked about an activity, or student surprise, level of fun or level of interest/boredom. The use of a survey or questionnaire was the most common means of assessing student motivation.

Twenty of the 26 papers that included student motivation in their research design used a survey or questionnaire; about half included open-ended questions on the survey, while the others used a Likert scale. Analysis was generally both quantitative and qualitative. The remaining papers did not give a clear description of the instrument used. In addition, nine studies used student interviews to assess motivation, while three studies used naturalistic observation. Student work also served as a source of data pertaining to student motivation, as three studies used this data collection method. Finally, four papers used teacher interviews or surveys.

All but one of the twenty-six studies that formally included a motivation measure reported positive motivation outcomes of some sort. Sixteen papers described a strong incidence of enhanced motivation, commenting that the majority of their study subjects found the instructional program (or other intervention) interesting, engaging, fun, etc. For example, Jones et al. (2006) reported, “All of the students who participated in the study reported that the instructional program was highly interesting with a mean rating of 8.65 on a scale of 1 (not interesting) to 10 (highly interesting)" (p. 349). Hoese and Nowicki (2001) included items related to student satisfaction with their biology course and found that the grade given for a student’s in-class presentation was positively related to the student’s satisfaction with the class.

Seven papers found mixed results with some motivation items showing positive gains, and others not. For example, Schaal and Bogner (2005) report, “ In a mid-lesson survey, the pupils in Group 1 scored significantly higher on their ‘well-being’ variable than those in Group 2 but felt significantly more bored during the lesson” (p. 34).

An additional 17 papers, while not formally including motivation measures, nevertheless included informal comments on student motivation, often coming from the teacher or from researcher observations. Authors used such comments to suggest areas of interest for possible future research, to help in interpreting the primary results, or otherwise providing insight into the learners' experience. Six of these papers reported lack of motivation (boredom, lack of ability to stay on task, etc.). For example, Windschitl (2001) reported, “Student lack of persistence was also a problem during the study. Few dyads were able to maintain their focus on the simulation exercises for the duration of the second week.” He recommended, “Future studies of this type should not neglect the role of motivation in students for these kinds of extended learning activities.” Hogan and Thomas (2001) also commented on lack of persistence, “When they [pairs of students] did have ideas about model revision [...] they did not show accompanying motivation to delve back into their models to make and explore the effects of changes.”

In some cases, the needs of the research interfered negatively with student motivation. Zohar and Nemet (2002) reported positive outcomes from their experiment group in the areas of content knowledge and argumentation skills. But they also commented, “…it is important to report that teachers and students who participated in this study generally were enthusiastic about the program. Nevertheless, they complained that repeating teaching and learning about moral dilemmas in 12 consecutive lessons was tiring.”

C. “Other” Student Outcomes

Thirty-three papers reported student outcomes not covered by the coding scheme. These included numerous themes such as: detailed analysis of the laboratory task; interaction with other students; affective measures; gender; and creativity.

The papers that included outcomes related to a more detailed analysis of the laboratory task reported on: the number of lessons/activities done by students and the time spent on task; the usability of laboratory software; the difficulty of the laboratory task; and the types of questions students asked during the activity. For example, Lidemann-Matthies (2005) studied how the number of lessons a teacher spent on the

‘Nature on the Way to School’ curriculum influenced students’ appreciation for native flora and fauna. The study found that “[t]eachers spent between 1 and 60 hours of lessons on the various activities of ‘Nature on the way to school.’ On average a class received 17 hours of lessons. The number of lessons taught had a strong effect on children’s preferences for both plants and animals. The more lessons a class received, the greater was the additional proportion of children that most of all appreciated one of the wild plants of Switzerland…” (Lidemann-Matthies, 2005, p. 663-664). Parr (2004) addressed the usability of customized software used during the laboratory activity. The paper reported that the majority of students where able to learn the software without much difficulty and that the icon-driven software interface helped young students enter data faster. (Parr, 2004, p. 237)

Papers that included outcomes related to student interaction with other students or other people reported on: peer collaboration and student-to-student interaction; the physical and behavioral impact of the activity; and student engagement toward the author of the paper. For example, Rueter (1999) reported that students using STELLA simulations during laboratory exercises “talked to each other and explained the facts and concepts in the material to each other. They then formulated an idea and tried it out on the computer simulation. […] The model exercise was very valuable in that it caused them to talk to each other” (p. 121). In reporting on the behavioral impacts of a residential field course, Amos and Reiss (2006) stated, “All teachers commented that pupils’ ‘behavior was good as, and often better than, at school.’ There were some minor behavior problems but most were dealt with quickly” (p. 41).

Some papers included affective measures such as: empathy and desire to help other species; the affinity for the garden site in which the activity was conducted; how context stimulates meaningful learning; and student attitudes towards things other than science (i.e. the environment or snakes). Tyson (2004) studied the effect of telic/paratelic tone of the instructor on the situational affinity and ecological understanding of fourth graders visiting the Red Butte Garden’s natural area. She found that students in both the telic and paratelic groups appeared to have an affinity for the natural area at the end of their two-hour experience. In looking at fifth-grade students’ attitudes towards snakes, Ford (1992) found that after participating in an informal program about snakes, “students’ attitudes toward snakes shifted to the more positive end of the scale. This short-term attitude shift mimicking the increase in knowledge was encouraging” (p. 146).

Finally, gender, creativity and vegetable eating habits were three other additional student outcomes reported by corpus papers. When evaluating a video-based frog dissection simulation, Akpan (1999) reported that “[n]o differences in posttest achievement or dissection performance were found between male and female participants in any condition” (p. 117). Teachers interviewed in Haigh (2007) commented that “I would carry out a similar [investigative practical work] programme in the future years because it is one of the few subjects at school where students can think for themselves and be creative in their ideas.” (p. 135). In looking at the impact of school gardening experiences on student vegetable consumption and understanding of ecoliteracy, Ratcliffe (2007) found that “gardening influenced factors associated with vegetable consumption, including increased variety eaten as measured by self-reported monthly consumption, and consumption of different vegetable varieties at school” (p. v).

11. GAP ANALYSIS AND LIMITATIONS OF THE STUDY

When we began this study, we expected to be able to identify gaps that might exist in the research literature. That is, we expected that the literature would provide insufficient evidence to enable generalizations about the contribution of labs or lab experiences to one or more important questions of interest to the field. While some of these gaps might be supplied by earlier research on, say, lab contributions to students' overcoming misconceptions, or gaining specific learning skills, it is also the case that educational research proceeds in such a way that questions can and should be revisited afresh. As research in the many sub-domains of science education, the findings in, say, cognitive psychology, classroom discourse studies, and biology itself, partially supercede earlier work. The 20-year window we adopted for this study has provided a lens on one era of research, in which the field has been shaped by the development and deployment of science education standards, misconception theory, and a growing interest in learning progressions in the sciences.

It seems safe to say that the research literature that we have examined, even if we had included more of the excluded papers, does not support firm conclusions about the lab component of the life sciences curriculum. Two key theoretical questions that lie behind our specific research questions for this study are:

[1] Does the research literature on biological lab experiences provide evidence about how to scaffold the development of biological reasoning over the course of a student's life sciences learning 'career" from middle-school through introductory college biology?

[2] If one takes the view that the labs in life science classes could constitute a curriculum across the grade levels, what evidence is there from the research literature of the completeness and developmental coherence of this course?

Examining the literature with these questions in mind, and through the lens of our research questions, it is the gaps that are most apparent. We address here some specific gaps which we find strategically important, and which could constitute important fields for further research and development.

a. We found insufficient research on special populations and how labs can be designed to support various kinds of learning challenges, whether individual or social in nature. For example, there were few papers addressing how gender interacts with the design or implementation of lab experiences. There were no papers that addressed possible interactions with SES or other demographic features which have been shown in educational research to (at least sometimes) play an important role in science learning, and only one paper addressing differentiated instruction.

b. The corpus provides little or no evidence about how lab experiences can contribute to the growth of biological reasoning over the grades, and in the growth of related investigative or inquiry skills, such as systematic observation or other qualitative skills, data analysis, experimental design, error analysis in biological investigations, or the handling of biological material. One can say that the field, as represented by these papers, does not even provide a skeletal framework for how these and other aspects of biological investigations should be addressed at different levels of accomplishment or age, and we saw no evidence that lab designs at the level of introductory college courses are any more sophisticated in the demands they make on their students than the labs in middle school.

c. Related to the prior point, only a few papers seem to describe and analyze labs that present the work in the context of science as a theory-driven enterprise, in which investigations are designed to address problems identified against a theoretical framework and problematic data. Thus, few labs seem to require the students to refer to even qualitative models of the phenomena under investigation, and even fewer expect students to build models (qualitative or quantitative) to understand the value of the lab. This is an important reason that lab activities are decontextualized both with regard to other labs in the course, and in the curriculum itself, a persistent problem with laboratory activities across the sciences (NRC 2006).

d. It is remarkable how few standard labs are examined in the literature, either studied in their own right, or as comparisons to innovative approaches which might be the focus of the specific study. The "innovation bias" of the literature means that the efficacy of a wide range of lab activities in addressing issues such as the needs of particular populations, subjects, or grade level understandings of specific concepts remains unassayed.

e. Again, there are significant areas of biology which are addressed sparsely or not at all, including many aspects of microbiology, anatomy, physiology, developmental biology, taxonomy. Even within the most -addressed areas, such as genetics and ecology, the coverage is extremely uneven.

f. Lab activities which are not specifically targeted on evolutionary questions do not explicitly engage the students with the biological distinctives, discipline-specific conceptualizations and methodologies characteristic of biological systems. Functional phenomena (e.g. in physiology or anatomy) are not placed in systemic context, either with regard to organismal/ecological implications or evolutionary dimensions. There is little evidence about how to scaffold the ability to study complex systems with high noise-to-signal ratio — the strategy in almost all cases (see above for discussion) is to work within a simplified system, with no exploration of how to relate such systems to more realistic ones.

g. The research literature provides very little evidence about the role of lab experiences in engaging student interest in life science topics, only a few papers exploring this issue.

h. Finally, we note that the value of the literature as a resource for further research or development — the translation of research into practice — is definitely limited by the generally poor description of the interventions being studied, as also noted by Minner et al. (2010) in their study of the impact of inquiry on student learning.

Limitations of the study, and notes on desirable collateral research

While this study is the first of its kind, and thus makes a significant contribution to the literature, we note that it might be improved upon in at least the following ways:

1. The time period being studied might be extended backwards. This would provide a fuller picture of what research may be available that bears on some of the gaps found in the current corpus.

2. The corpus could be expanded by including more journals, more thesis databases, and literature in more languages than English. While we believe that we have consulted the most important journals in the field, there are probably others to be searched, as well as journals in related fields such as educational and cognitive psychology.

3. We have not done a systematic analysis with respect to the gaps noted above of the papers excluded from this study. As we continue to analyze our data for publication, we intend to revisit these papers with this aim in mind.

Finally, we note that our focus on the research literature automatically excluded insights relating to our research questions that could be gleaned from an analysis of lab manuals and texts; classroom observations; and the very large number of articles on innovation and problem solving in biology teaching practice. All three of these areas merit examination in more detail, and each can contribute valuably to further research. In particular, the "practitioner literature" appears to be an important source of hypotheses for further study, since the journals devoted to biology teaching include reports on classroom practice from across the grade levels, and across many domains of biology, providing better topical representation than the research literature does. It is likely that an analysis of some of this literature, using some version of the analytic framework we have developed in this project, should be our next research project.

REFERENCES

AAAS. (2001). Project 2061: Atlas of Scientific Literacy. Washington DC: AAAS.

Abell and Lederman (2007)

Bell, P. (2004). *The school science laboratory: Considerations of learning, technology, and scientific practice.* Paper prepared for the Committee onHigh School Science Laboratories: Role and Vision. http://www7.nationalacademies.org/bose/July_12_13_2004_high_school_labs_ meeting_ agenda.html, downloaded May 2 2006.

Choppin, B., and R. Frankel (1976). The three most interesting things. *Studies in Educational Evaluation* 2: 57-61.

Deboer, G.E. (1991) *A history of ideas in science education: implications for practice*. New York: Teachers College Press.

Driver, R., J. Leach, R. Millar, and P. Scott (1996) *Young people's images of science*. Philadelphia: Open University Press.

Duckworth 1991

Janovy, J., Jr. (2003) *Teaching in Eden: Lessons from Cedar Point*. New York: Routledge Falmer.

Lazarowitz, R. and P. Tamir (1994) Research on using laboratory instruction in science. In Gabel, D.L., ed., *Handbook of research on science teaching and learning*. New York: Macmillan Publishing Company.

Lurie, E. (1988) *Louis Agassiz: A life in science*. Baltimore: The Johns Hopkins University Press.

Mayr, E. (2004) *What makes biology unique? Considerations on the autonomy of a scientific discipline.* Cambridge: Cambridge University Press.

Mayr, E. (1982). *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Cambridge, MA: Belknap Press of Harvard University Press.

NAEP, (1999). Trends in Academic Progress.

National Research Council. (2007) *Taking science to school: Learning and teaching science in grades K-8*. Committee on Science Learning, Kindergarten through Eighth Grade. Duschl, R. A., H.A. Schweingruber, and A.W. Shouse, editors. Washington DC: The National Academies Press.

National Research Council. (2006). *America's Lab Report: Investigations in High School Science.* Committee of High School Science Laboratories: Role and Vision, S.R. Singer, M.L. Hilton, & H.A. Schweingruber (eds.) Board on Science Education, Center for Education. Division of Behavioral and Social Sciences and Education. Washington DC: The National Academies Press.

National Research Council. (1996). *National Science Education Standards*. Washington DC: The National Academies Press.

National Science Board (2006) *Science and Engineering Indicators 2006*. Two vols. Arlington, VA: National Science Foundation.

Passmore, C. & Stewart, J. (2002). A modeling approach to teaching evolutionary biology in high school. *Journal of Research in Science Teaching, 39, 185-204.*

Rudolph, J. L. & Stewart, J. H. (1998). Evolution and the nature of science: On the historical discord and its implications for education. *Journal of Research in Science Teaching,* 35, 1069-1089.

Stratford et al. 1998

Sterelny, K. & Griffiths, P.E. (1999). *Sex and Death: An introduction to the philosophy of biology.* Chicago IL: University of Chicago Press.

Trumper, R. (2006) Factors affecting junior high school students' interest in biology. *Science Education International*: 17 (1): 31-48.

UNDER THE MICROSCOPE CORPUS

Abell, S. K. & M. Roth (1995). Reflection on a fifth-grade life science lesson: making sense of children's understanding of scientific models. *International Journal of Science Education*,17, 59-74.

Aho, L., Hupio, J. & Huttunen, S. (1993). Learning science by practical work in Finnish primary school using materials familiar from the environment: a pilot study. *International Journal of Science Education,*15, 497-507.

Ajewole, G. A. (1991). Effects of discovery and expository instructional methods on the attitude of students to biology. *Journal of Research in Science Teaching,* 28, 401-409.

Akpan, J. P. & T. Andre (1999). The effect of prior dissection simulation on middle school students' dissection performance and understanding of the anatomy and morphology of the frog. *Journal of Science Education and Technology*, 8, 107-121.

Alsop, S. (2001). Seeking emotional involvement in science education: food-chains and webs. *School Science Review,* 83, 63-68.

Amos, R. & M. Reiss (2006). What contribution can residential field courses make to the education of 11-14 year-olds? *School Science Review,* 88,37-44.

Ashburn, S. J., Eichinger, D.C., Witham, S.A., Cross, V.D., Krockover, G.H., Pae, T-I., Islam, S. & Robinson, J.P. (2002). The bioscope initiative: Integrating technology into the biology classroom. *American Biology Teacher,* 64, 503-510.

Bailey, S. & R. Watson (1998). Establishing basic ecological understanding in younger pupils: a pilot evaluation of a strategy based on drama/role play. *International Journal of Science Education,* 20, 139-152.

Banet, E. & E. Ayuso (2000). Teaching genetics at secondary school: A strategy for teaching about the location of inheritance information. *Science Education,* 84, 313-351.

Banet, E. & G. E. Ayuso (2003). Teaching of biological inheritance and evolution of living beings in secondary school. *International Journal of Science Education,* 25, 373.

Banet, E. & F. Nunez (1997). Teaching and learning about human nutrition: a constructivist approach. *International Journal of Science Education,* 19, 1169 - 1194.

Beardsley, P. M. (2004). Middle school student learning in evolution: Are current standards achievable? *American Biology Teacher,* 66, 604-612.

Bowker, R. & A. Jasper (2007). 'Don't forget your leech socks'! Children's learning during an Eden Education Officer's workshop. *Research in Science and Technological Education,* 25, 135-150.

Brisbin, T. E. (2000). *An Account of Novice Phylogenetic Tree Construction from the Problem-Solving Research Tradition.* Doctor of Philosophy, Western Michigan

University.

Browning, M. E. & J. D. Lehman (1988). Identification of student misconceptions in genetics problem solving via computer program. *Journal of Research in Science Teaching,* 25, 747-761.

Buckley, B. C. (2000). Interactive multimedia and model-based learning in biology. *International Journal of Science Education*, 22, 895-935.

Charney, J., Hmelo-Silver, C.E., Sofer, W., Neigeborn, L., Coletta, S. & Nemeroff, M. (2007). Cognitive apprenticeship in science through immersion in laboratory practices. *International Journal of Science Education,* 29, 195-213.

Chinnici, J. P., J. W. Yue, et al. (2004). Students as "human chromosomes" in role-playing mitosis and meiosis. *American Biology Teacher, 66, 35-39.*

Christianson, R. G. & K. M. Fisher (1999). Comparison of student learning about diffusion and osmosis in constructivist and traditional classrooms. *International Journal of Science Education,* 21, 687-698.

Cross, T. R. & V. E. Cross (2004). Scalpel or mouse? A statistical comparison of real and virtual frog dissections. *American Biology Teacher,* 66, 408-411.

Eaton, D. (1998). *Cognitive and affective learning in outdoor education.* Doctor of Education, University of Toronto.

Ergazaki, M., Komis, V. & Zogza, V. (2005). High-school students' reasoning while constructing plant growth models in a computer-supported educational environment. *International Journal of Science Education,* 27, 909-933.

Evans, S., S. Dixon, et al. (2006). Pupils' knowledge of birds: how good is it and where does it come from? *School Science Review,* 88, 93-98.

Faryniarz, J. V. & L. G. Lockwood (1992). Effectiveness of microcomputer simulations in stimulating environmental problem solving by community college students. Journal of Research in Science Teaching, 29, 453-470.

Fedak, J., W. Belzer, et al. (1990). Videomicroscopy and improved student attitudes. *American Biology Teacher,* 52, 419-421.

Ford, C. S. (1992). *The influence of experiential nonformal learning strategies on fifth-grade students' knowledge and attitudes toward snakes.* Doctor of Philosophy, Kansas State University.

Gallucci, K. K. (2007). *The case method of instruction, conceptual change, and student attitude.* Doctor of Philosophy, North Carolina State University.

Gelbart, H. & A. Yarden (2006). Learning genetics through an authentic research simulation in bioinformatics. *Journal of Biological Education,* 40, 107-112.

Geraedts, C. L. & K. T. Boersma (2006). Reinventing natural selection. *International Journal of Science Education,*28, 843-870.

Gilman, S. L. (2006). Do online labs work? An assessment of an online lab on cell division. *American Biology Teacher Online Publication*: 131-134.

Grace, M. M. & M. Ratcliffe (2002). The science and values that young people draw upon to make decisions about biological conservation issues. *International Journal of Science Education,* 24, 1157-1169.

Hafner, R. & J. Stewart (1995). Revising explanatory models to accommodate anomalous genetic phenomena: Problem solving in the "context of discovery". *Science Education,* 79, 111-146.

Haigh, M. A. (2007). Can investigative practical work in high school biology foster creativity? *Research in Science Education* 37: 123-140.

Hall, D. A. & D. W. McCurdy (1990). A comparison of a Biological Sciences Curriculum Study (BSCS) laboratory and a traditional laboratory at two private liberal arts colleges. *Journal of Research in Science Teaching,* 27, 625-636.

Hamilton-Ekeke, J.-T. (2007). Relative effectiveness of expository and field trip methods of teaching on students' achievement in ecology. *International Journal of Science Education,* 29, 1869-1889.

Helms, J. V. (1998). Science and/in the community: context and goals in practical work. *International Journal of Science Education,* 20, 643-653.

Hickey, D. T., Kindfield, A.C.H., Horwitz, P., & Christie, M.A.T. (2003). Integrating curriculum, instruction, assessment, and evaluation in a technology-supported genetics learning environment. *American Educational Research Journal,* 40, 495-538.

Hoese, W. J. & S. Nowicki (2001). Using "the organism" as a conceptual focus in an introductory biology course. *American Biology Teacher,* 63, 176-182.

Hogan, K. (2002). Small groups' ecological reasoning while making an environmental management decision. *Journal of Research in Science Teaching,* 29, 341-368.

Hogan, K. & D. Thomas (2001). Cognitive comparisons of students' systems modeling in ecology. *Journal of Science Education and Technology,* 10, 319-345.

Hounshell, P. B. & J. Stanford R. Hill (1989). The microcomputer and achievement and attitudes in high school biology. *Journal of Research in Science Teaching,* 26, 543-549.

Huang, S. D. & J. Aloi (1991). The impact of using interactive video in teaching general biology. *American Biology Teacher,* 53, 281-284.

Huppert, J., Lomask, S.M. & Lazarowitz, R. (2002). Computer simulations in the high school: students' cognitive stages, science process skills and academic achievement in microbiology. *International Journal of Science Education,* 24, 803-821.

Huppert, J., Yaakobi, J. & Lazrowitz, R. (1998). Learning Microbiology with Computer Simulations: students' academic achievement by method and gender. *Research in Science and Technological Education,* 17, 231-245.

Jimenez-Aleixandre, M-P. & Pereira-Munoz, C. (2002). Knowledge producers or knowledge consumers? Argumentation and decision making about environmental management. *International Journal of Science Education,* 24, 1171-1190.

Johnson, M. A. & A. E. Lawson (1998). What are the relative effects of reasoning ability and prior knowledge on biology achievement in expository and inquiry classes. *Journal of Research in Science Teaching,* 35, 89-103.

Jones, M. G., T. Andre, et al. (2003). Learning at the nanoscale: The impact of students' use of remote microscopy on concepts of viruses, scale, and microscopy. *Journal of Research in Science Teaching,* 40, 303-322.

Jones, M. G., J. Minogue, et al. (2006). Visualizing without vision at the microscale: Students with visual impairments explore cells with touch. *Journal of Science Education and Technology,* 15, 345-351.

Kiboss, J. K., Ndirangu, M. &Wekesa, E.W. (2004). Effectiveness of a computer-mediated simulations program in school biology on pupils' learning outcomes in cell theory. *Journal of Science Education and Technology,* 13, 207-213.

Kinzie, M. B. (1993). The effects of an interactive dissection simulation on the performance and achievement of high school biology students. *Journal of Research in Science Teaching,* 30, 989-1000.

Knapp, D. & E. Barrie (2001). Content evaluation of an environmental science field trip. *Journal of Science Education and Technology,* 10, 351-357.

Koomen, M. J. H. (2006). *Listening to their voices: The essence of the experience of special and regular education students as they learn monarch, Danaus plexippus, biology and ecology.* Doctor of Philosophy, University of Minnesota.

LaFlamme, M. (2004). *Motivating conservation: Learning to care for other species in a local ecological community.* Doctor of Philosophy in Educational Thought and Sociocultural Studies, University of New Mexico.

Lavoie, D. R. (1999). Effects of emphasizing hypothetico-predictive reasoning within the science learning cycle on high school student's process skills and conceptual understandings in biology. *Journal of Research in Science Teaching,* 36, 1127-1147.

Lavoie, D. R. & Good, R. (1988). The nature and use of prediction skills in a biological computer simulation. *Journal of Research in Science Teaching,* 25, 335-360.

Law, N. & Y. Lee (2004). Using an iconic modelling tool to support the learning of genetics concepts. *Journal of Biological Education,* 38, 126-133.

Leonard, W. H. (1992). A comparison of student performance following instruction by interactive videodisc versus conventional laboratory. *Journal of Research in Science Teaching,* 29, 93-102.

Lewis, J. & J. Leach (2006). Discussion of Socio-scientific Issues: The role of science knowledge. *International Journal of Science Education,* 28, 1267-1287.

Lightburn, M. E. & B. J. Fraser (2007). Classroom environment and student outcomes among students using anthropometry activities in high-school science. *Research in Science and Technological Education,* 25, 153-166.

Lindemann-Matthies, P. (2005). 'Loveable' mammals and 'lifeless' plants: how children's interest in common local organisms can be enhanced through observation of nature. *International Journal of Science Education,* 27, 655-677.

Lindemann-Matthies, P. (2006). Investigating Nature on the Way to School: Responses to an educational programme by teachers and their pupils. *International Journal of Science Education,* 28, 895-918.

Lumpe, A. T. (1995). Peer collaboration and concept development: Learning about photosynthesis. *Journal of Research in Science Teaching,* 32, 71-98.

Magntorn, O. & G. Hellden (2007). Reading New Environments: Students' ability to generalise their understanding between different ecosystems. *International*

Journal of Science Education, 29, 67-100.

Magntorn, O. & G. Helldén (2007). Reading nature from a 'bottom-up' perspective. *Journal of Biological Education,* 41, 68-75.

Manzanal, R. F., Barreiro, L.M.R. & Jimenez, M.C. (1999). Relationship between ecology fieldwork and student attitudes toward environmental protection. *Journal of Research in Science Teaching,* 36, 431-453.

Marbach-Ad, G. & L. A. Claassen (2001). Improving students' questions in inquiry labs. *American Biology Teacher,* 63, 410-419.

Marbach-Ad, G., Rotbain, Y. & Stavy, R. (2005). Using a bead model to teach A-level molecular biology. *School Science Review,* 87, 39-52.

Nehm, R. H. & A. F. Budd (2006). Missing "links" in bioinformatics education: Expanding students' conceptions of bioinformatics using a biodiversity database of living and fossil reef corals. *American Biology Teacher Online Publication*: 91-97.

Olsher, G., Berl, D.B. & Dreyfus, A. (1999). Biotechnologies as a context for enhancing junior high-school students' ability to ask meaningful questions about abstract biological processes. *International Journal of Science Education,* 21, 137-153.

Parr, C. S., Jones, T. & Songer, N. (2004). Evaluation of a handheld data collection interface for science learning. *Journal of Science Education and Technology,* 13, 233-242.

Passmore, C. & J. Stewart (2002). A modeling approach to teaching evolutionary biology in high schools. *Journal of Research in Science Teaching,* 39, 185-204.

Pata, K. & T. Sarapuu (2006). A comparison of reasoning processes in a collaborative modelling environment: Learning about genetics problems using virtual chat. *International Journal of Science Education*, 28, 1347-1368.

Prokop, P., Tuncer, G. & Kvasnicak, R. (2007). Short-term effects of field programme on students' knowledge and attitude toward Biology: A Slovak experience. *Journal of Science Education and Technology,* 16, 247-255.

Rahm, J. (2002). Emergent learning opportunities in an inner-city youth gardening program. *Journal of Research in Science Teaching,* 39, 164-184.

Randler, C. & F. X. Bogner (2006). Cognitive achievements in identification skills. *Journal of Biological Education,* 40, 161-165.

Randler, C. & M. Hulde (2007). Hands-on versus teacher-centered experiments in soil ecology. *Research in Science and Technological Education,* 25, 329-338.

Ratanapojnard, S. (2001). *Community-oriented biodiversity environmental education: Its effect on knowledge, values, and behavior among rural fifth- and sixth-grade students in Northeastern Thailand.* Doctor of Philosophy, Yale University.

Ratcliffe, M. M. (2007). *Garden-based education in school settings: The effects on children's vegetable consumption, vegetable preferences and ecoliteracy.* Doctor of Philosophy in the Program in Agriculture, Food and the Environment, Tufts Universtiy.

Reuter, J. J. (2005). *Using the BiodatamationTM strategy to learn introductory college Biology: Value-added effects on selected students' conceptual understanding and* *conceptual integration of the processes of photosynthesis and cellular respiration.* Doctor of Philosophy, Tulane University.

Rotbain, Y., Marbach-Ad, G. & Stavy, R. (2006). Effect of bead and illustrations models on high school students' achievement in molecular genetics. *Journal of Research in Science Teaching,* 43, 500-529.

Rothhaar, R., Pittendrigh, B.R. &Orvis, K.S. (2006). The LEGO analogy model for teaching gene sequencing and biotechnology. *Journal of Biological Education,* 40, 166-171.

Rueter, J. G. & N. A. Perrin (1999). Using a web simulation to teach food web dynamics. *American Biology Teacher,* 61, 116-123.

Sanger, M. J., D. M. Brecheisen, et al. (2001). Can computer animations affect college biology students' conceptions about diffusion and osmosis? *American Biology Teacher,* 63, 104-109.

Schaal, S. & F. X. Bogner (2005). Human visual perception — learning at workstations. *Journal of Biological Education,* 40, 32-37.

Seethaler, S. & M. Linn (2004). Genetically modified food in perspective: an inquiry-based curriculum to help middle school students make sense of tradeoffs. *International Journal of Science Education*, 26, 1765-1785.

Sewell, R. D. E. & R. G. Stevens (1995). Multimedia computer technology as a tool for teaching and assessment of biological science. *Journal of Biological Education,* 29, 27.

Slack, S. J. & J. Stewart (1990). High school students' problem-solving performance on realistic genetics problems. *Journal of Research in Science Teaching,* 27, 55-67.

Soyibo, K. & A. Hudson (2000). Effects of Computer-assisted Instruction (CAI) on 11th Graders' Attitudes to Biology and CAI and Understanding of Reproduction in Plants and Animals. *Research in Science and Technological Education,* 18, 191-199.

Stratford, S. J., Krajcik, J. & Soloway, E. (1998). Secondary student's dynamic modeling processes: Analyzing, reasoning about, synthesizing, and testing models of stream ecosystems. *Journal of Science Education and Technology,* 7, 215-234.

Strauss, R. T. & M. B. Kinzie (1994). Student achievement and attitudes in a pilot study comparing an interactive videodisc simulation to conventional dissection. *American Biology Teacher,* 56, 398-402.

Talsma, V. L. (2004). *Student scientific understandings in a ninth grade project-based science classroom: A river runs through it.* Doctor of Philosophy in Education, University of Michigan.

Tamir, P. & Y. Shcurr (1997). Back to living animals: An extracurricular course for fifth-grade pupils. *Journal of Biological Education,* 31, 300-305.

Taraban, R., Box, C., Myers, R., Pollard, R. & Bowen, C.W. (2007). Effects of active-learning experiences on achievement, attitudes, and behaviors in high school biology. *Journal of Research in Science Teaching,* 44, 960-979.

Tomkins, S. P. & S. D. Tunnicliffe (2001). Looking for ideas: Observation, interpretation and hypothesis-making by 12-year-old pupils undertaking science investigations.

International Journal of Science Education, 23, 791-813.

Tregidgo, D. & M. Ratcliffe (2000). The use of modelling for improving pupils' learning about cells. *School Science Review,* 81, 53-58.

Tsui, C.-Y. & D. A. Treagust (2004). Conceptual change in learning genetics: An ontological perspective. *Research in Science and Technological Education,* 22, 185-202.

Tsui, C.-y. & D. F. Treagust (2004).Motivational Aspects of Learning Genetics with Interactive Multimedia. *American Biology Teacher,* 66, 277-285.

Tyson, B. E. (2004). *Fostering emotional affinity and ecological understanding in environmental education experience: An application of reversal theory and generative teaching method.* Doctor of Philosophy, University of Utah.

Verhoeff, R. P., Waarlo, A.J. & Boersma, K.T. (2008). Systems modelling and the development of coherent understanding of cell biology. *International Journal of Science Education,* 30 543-568.

Webb, P. & G. Boltt (1991). High school pupils' and first-year university students' responses to questions based on data. *Journal of Biological Education* 25, 119, 14pp.

Windschitl, M. (2001). Using simulations in the middle school: Does assertiveness of dyad partners influence conceptual change? *International Journal of Science Education,* 23, 17-32.

Windschitl, M. & T. Andre (1998). Using computer simulations to enhance conceptual change: the roles of constructivist instruction and student epistemological beliefs. *Journal of Research in Science Teaching,* 35, 145-160.

Witenoff, S. & R. Lazarowitz (1993). Restructuring laboratory worksheets for junior high school biology students in the heterogeneous classroom. *Research in Science and Technological Education,* 11, 225-239.

Wu, Y.-T. & C.-C. Tsai (2005). Effects of constructivist-oriented instruction on elementary school students' cognitive structures. *Journal of Biological Education,* 39, 113-119.

Wynne, C. F., Stewart, J. & Passmore, C. (2001). High school students' use of meiosis when solving genetics problems. *International Journal of Science Education,* 23, 501-515.

Yoon, S. A. (2007). An evolutionary approach to harnessing complex systems thinking in the science and technology classroom. *International Journal of Science Education,* 30, 1-32.

Zohar, A. (1996). Transfer and retention of reasoning strategies taught in biological contexts. *Research in Science and Technological Education,* 14, 205-219.

Zohar, A. (1998). Result or conclusion? Students' differentation between experimental results and conclusions. *Journal of Biological Education,* 32, 53, 7pp.

Zohar, A. & F. Nemet (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching,* 39, 35-62.

Zoldosova, K. & P. Prokop (2006). Education in the field influences children's ideas and interest towards science. *Journal of Science Education and Technology,* 15, 304-313.

APPENDIX 1 JOURNALS SEARCHED

Alberta Journal of Educational research
American Education Research Journal
Australian Journal of Education
Bioscene
British Educational Research Journal
Canadian Journal of Education
Cognition and Instruction
Elementary school journal
International Journal of Science and Mathematics Education
International Journal of Science Education
Journal of Biological Education
Journal of College Biology Teaching
Journal of College Science Teaching
Journal of Research on Science Teaching
Research in Science Education
Research in Science Education
School Science Review
Science and Education
Science Education
Studies in Science education

---

[1] We use the definition used by the NRC (2006): Laboratory experiences provide opportunities for students to interact directly with the material world (or with data drawn from the material world), using tools, data collection techniques, models, and theories of science. They may be separate "labs," or form one component of "integrated instructional units." (See NRC 2006, pg. 78ff.)

[2] This work supported by NSF/REESE grant 08-15190