Letters to the Editor

Re: The Use of a Knowledge Survey as an Indicator of Student Learning in an Introductory Biology Course

Edward B. Nuhfer* and Delores Knipp†

*Department of Geology and Center for Teaching and Learning, Idaho State University, Pocatello, ID 83209; and †Department of Physics, United States Air Force Academy, United States Air Force Academy, CO 80840

Reliability is a fundamental quality of the internal consistency of any measuring instrument. Researchers sometimes use “reliable” and its derivative terms, yet fail to address reliability. Bowers et al. (2005) claimed that the knowledge survey (KS) “does not reliably measure student learning as measured by final grades or exam questions.” They addressed their purpose (p. 311) “to evaluate how closely students’ performance track with their confidence in their knowledge of the course material,” through correlating “plotted pre- and post-KS scores against final grades (p. 314).” This approach assumes that tests/grades of unknown reliability are appropriate standards for judging other measures. Their article offers a case study in drawing conclusions without considering reliability.

The split-halves Spearman-Brown reliability (R) measure (Jacobs and Chase, 1992) derives from the r (r) obtained from individuals’ scores on two halves of a single test. It is a routine method for quantifying reliability and is applicable to both tests and knowledge surveys. When Bowers et al. attributed specific claims to us: “They report that KS results represent changes in students’ learning (p. 311),” and “... are a good representation of student knowledge (p. 316),” they omitted mention of the surprising reliability that characterizes knowledge surveys’ pre- and postcourse measures (Figure 1).

Figure 2 shows a contrasting split-halves reliability for a faculty-made test of unusually high reliability that yields an r of r = 0.8 and a reliability coefficient of R = 0.9. Good faculty-made tests achieve R >0.6 (Jacobs and Chase, 1992), but many yield reliability coefficients of <0.3 (Raoul Arreola, personal communication, April 2005; Theall et al., 2005). In correlating the test in Figure 2 with its equivalent KS, the maximum correlation to expect would be about r = 0.8 (the degree to which the less reliable instrument can correlate with itself). We believe r should be still lower, because we developed knowledge surveys to sample cognitive/affective domains that only partially overlap those sampled by tests (Wirth et al., 2005).

Next, consider the reliability of grades derived by averaging several tests. An example appears in Figure 3. Suppose some tests lack significant correlation with other tests, but, as is common practice, we average all the tests anyway to produce grades. The reliability of grades derived from averaging a mix of tests should be lower than the most reliable tests used during the course. If we correlate grades and KS ratings, what, then, are we actually correlating? Without reliability measures, we cannot know.

Short-answer tests longer than 40 items usually have good reliability. Such tests define achievement based largely on knowing, as manifested through short-answer test-taking skills under timed conditions. Tests that address open-ended challenges sample other cognitive/affective domains to define achievement based largely on doing, as manifested through written reports and products generated after discussion, reflection, and revision. Such tests serve as much to promote learning and to mentor students to high-level thinking as to produce grades. High reliability is difficult to achieve using open-ended challenges, and quantifying it requires methods too cumbersome for routine classroom use. Many such tests within Figure 3 contribute to lower overall grade reliability than if achievement were derived solely from short-answer tests. One should know general reliability of one’s tests/grades to understand what comparisons are possible, but it is less important to optimize a correlation between tests and knowledge surveys than it is to mentor students to engage open-ended challenges. Exceptionally effective courses may show near zero reliability in plots such as Figure 3. In such cases, paired comparisons are impossible, but the assessment value of a well-designed KS as a reliable, complementary measure becomes apparent.

Unrealistic expectations for high numerical correlation coefficients between grades and knowledge surveys persist until one understands limits imposed by reliability. Thereafter, understanding permits better interpretations. The positive numerical correlations Bowers et al. (2005) reported as “low” between their postcourse knowledge surveys and grades then seem surprisingly high, given limits imposed by reliability of tests and grades.

We emphasized assessment of learning in classes through use of aggregate data (Nuhfer and Knipp, 2003), which differs from Bowers et al.’s evaluative efforts to predict an individual’s grades from her/his knowledge surveys. At the class level, we were most impressed by Bowers et al.’s Figure 1. It revealed the pattern change from no correlation be-
between grades and the precourse KS to a persistent positive correlation between grades and the postcourse KS. Their tools (tests and the knowledge survey) remained constant through their pre- and postcorrelations, so the profound pattern change seen in every class seems most simply explained by students’ increased understanding of specific content. We suggest correlating the high-reliability pre- and postcourse KS measures as an additional change indicator.

Although Bowers et al.’s article derived from our work, neither the authors from Nuhfer’s own Idaho State University campus nor the Journal’s editors engaged us in review. The result is an article including several attributions to knowledge surveys and our thoughts/intentions that we disclaim. Readers who compare our ideas about knowledge surveys as presented by Bowers et al. with those published in our words (Nuhfer and Knipp, 2003; Theall et al., 2005; Wirth et al. 2005) should anticipate discrepancies.

REFERENCES


RESPONSE: Re: The Use of a Knowledge Survey as an Indicator of Student Learning in an Introductory Biology Course

Nancy Bowers* and Maureen Brandon†

*Instructional Design and Technology Solutions, Potsdam, NY 13676; and †Idaho State University, Pocatello, ID 83209

We appreciate Nuhfer and Knipp's comments on our article, and we acknowledge that there are multiple ways to analyze the effectiveness of knowledge surveys as they relate to student learning. Because of the paucity of peer-reviewed research concerning the implementation and interpretation of knowledge surveys, there is no generally accepted method for analysis. The statistical methods we used to address our question of the relationship between confidence levels (e.g., knowledge survey results) and student performance are sound and appropriate. We encourage others, including Nuhfer and Knipp, to publish their findings through the peer-review process, and we look forward to reading those articles.

RESPONSE: Re: The Use of a Knowledge Survey as an Indicator of Student Learning in an Introductory Biology Course

Diane Ebert-May and Everett P. Weber

Department of Plant Biology, Michigan State University, East Lansing, MI 48824

What is the correlation between students' confidence in their level of knowledge and comprehension in a course and their actual performance, as judged by their grades? The research by Bowers, Brandon, and Hill (Bowers et al., 2005) investigated this question using knowledge surveys developed by Nuhfer and Knipp (2003) to determine students' perceived self-efficacy (belief in one's capability to carry out an action successfully). A knowledge survey (KS) is a series of content-based questions on topics presented in the course. Students are not asked to answer the questions, but merely to indicate for each question their confidence that they could answer it correctly. Evidence from Bowers et al. did not support the claim by Nuhfer and Knipp that students' learning can be predicted by perceived self-efficacy levels; rather, their data indicated that the correlation between student confidence and final grades is negligible. In response to this finding, Nuhfer and Knipp challenged the authors' interpretation of their results (see above letter). We are writing to comment on this controversy and to address a more fundamental question: How useful are knowledge surveys as assessment tools?

DOI: 10.1187/cbe.06–07–0173
Address correspondence to: Maureen Brandon (branmaur@isu.edu).

DOI: 10.1187/cbe.06–07–0174
Address correspondence to: Diane Ebert-May (ebertmay@msu.edu).
indeed be a valid link between the KS scores and student understanding reported by Bowers et al.

Reliability and validity are both critical aspects of assessment. Differences in interpretation of assessment results occur and merit productive debate. As Nuhfer and Knipp point out, assessments that define achievement based on lower-level thinking (e.g., knowledge and comprehension) usually have high reliability. If the goal of instruction is for students to demonstrate gains in their knowledge and comprehension of the subject, the KS may be useful. However, science involves more than mastering facts. The KS is not designed to probe students’ confidence about their ability to actively engage in processes of science or higher-level thinking such as analysis or synthesis. As we attempt to assess students’ critical thinking abilities using open-ended problems, determining reliability and validity becomes more difficult (Batzli et al., 2006). Ultimately, we wish to ascertain if assessments reliably measure what we want students to know and be able to do based on the goals and objectives of instruction. Are students becoming more sophisticated in their ability to solve complex problems that do not have single answers, high degrees of completeness, certainty or correctness? Is our instruction providing students guidance and practice in doing so?

Both papers agree that perceived self-efficacy is a key attribute when learning difficult subjects and addressing complex tasks. According to theory, as self-efficacy increases, students are more willing to undertake more complex tasks and think more about complex ideas and problems (Baldwin et al., 1999). Instruction that nurtures critical thinking enables students to gain deeper understanding of content knowledge. Specific guiding questions can facilitate students’ engagement with a problem and draw upon their prior knowledge or identify what they do not know (Ebert-May et al., 2006). Pretests that identify students’ comprehension or misconceptions about a topic can be used to guide instruction.

The peer-reviewed literature about knowledge surveys is sparse, and the relationship of perceived self-efficacy to performance merits further research. However, if critical thinking is the ultimate goal, a KS is unlikely to be a useful assessment. The potential value of knowledge surveys thus revolves around the question of whether “covering” content by the instructor is more important in an undergraduate science course than students “uncovering” content through problem solving and critical thinking. Whether instructors choose to use the KS should depend on their student learning goals. Perhaps instructors need to increase self-confidence in their own ability to promote higher-level thinking by their students.

REFERENCES


Reliability is a fundamental quality of the internal consistency of any measuring instrument. Researchers sometimes use “reliable” and its derivative terms, yet fail to address reliability. Bowers et al. (2005) claimed that the knowledge survey (KS) “does not reliably measure student learning as measured by final grades or exam questions.” They addressed their purpose (p. 311) “to evaluate how closely students’ performance track with their confidence in their knowledge of the course material,” through correlating “plotted pre- and post-KS scores against final grades (p. 314).” This approach assumes that tests/grades of unknown reliability are appropriate standards for judging other measures. Their article offers a case study in drawing conclusions without considering reliability.

The split-halves Spearman-Brown reliability (R) measure (Jacobs and Chase, 1992) derives from the r (r) obtained from individuals’ scores on two halves of a single test. It is a routine method for quantifying reliability and is applicable to both tests and knowledge surveys. When Bowers et al. attributed specific claims to us: “They report that KS results represent changes in students’ learning (p. 311),” and “. . . are a good representation of student knowledge (p. 316),” they omitted mention of the surprising reliability that characterizes knowledge surveys’ pre- and postcourse measures (Figure 1).

Figure 2 shows a contrasting split-halves reliability for a faculty-made test of unusually high reliability that yields an r of r = 0.8 and a reliability coefficient of R = 0.9. Good faculty-made tests achieve R > 0.6 (Jacobs and Chase, 1992), but many yield reliability coefficients of < 0.3 (Raoul Arreola, personal communication, April 2005; Theall et al., 2005). In correlating the test in Figure 2 with its equivalent KS, the maximum correlation to expect would be about r = 0.8 (the degree to which the less reliable instrument can correlate with itself). We believe r should be still lower, because we developed knowledge surveys to sample cognitive/affective domains that only partially overlap those sampled by tests (Wirth et al., 2005).

Next, consider the reliability of grades derived by averaging several tests. An example appears in Figure 3. Suppose some tests lack significant correlation with other tests, but, as is common practice, we average all the tests anyway to produce grades. The reliability of grades derived from averaging a mix of tests should be lower than the most reliable tests used during the course. If we correlate grades and KS ratings, what, then, are we actually correlating? Without reliability measures, we cannot know.

Short-answer tests longer than 40 items usually have good reliability. Such tests define achievement based largely on knowing, as manifested through short-answer test-taking skills under timed conditions. Tests that address open-ended challenges sample other cognitive/affective domains to define achievement based largely on doing, as manifested through written reports and products generated after discussion, reflection, and revision. Such tests serve as much to promote learning and to mentor students to high-level thinking as to produce grades. High reliability is difficult to achieve using open-ended challenges, and quantifying it requires methods too cumbersome for routine classroom use. Many such tests within Figure 3 contribute to lower overall grade reliability than if achievement were derived solely from short-answer tests. One should know general reliability of one’s tests/grades to understand what comparisons are possible, but it is less important to optimize a correlation between tests and knowledge surveys than it is to mentor students to engage open-ended challenges. Exceptionally effective courses may show near zero reliability in plots such as Figure 3. In such cases, paired comparisons are impossible, but the assessment value of a well-designed KS as a reliable, complementary measure becomes apparent.

Unrealistic expectations for high numerical correlation coefficients between grades and knowledge surveys persist until one understands limits imposed by reliability. Thereafter, understanding permits better interpretations. The positive numerical correlations Bowers et al. (2005) reported as “low” between their postcourse knowledge surveys and grades then seem surprisingly high, given limits imposed by reliability of tests and grades.

We emphasized assessment of learning in classes through use of aggregate data (Nuhfer and Knipp, 2003), which differs from Bowers et al.’s evaluative efforts to predict an individual’s grades from her/his knowledge surveys. At the class level, we were most impressed by Bowers et al.’s Figure 1. It revealed the pattern change from no correlation be-
between grades and the precourse KS to a persistent positive correlation between grades and the postcourse KS. Their tools (tests and the knowledge survey) remained constant through their pre- and postcorrelations, so the profound pattern change seen in every class seems most simply explained by students’ increased understanding of specific content. We suggest correlating the high-reliability pre- and postcourse KS measures as an additional change indicator.

Although Bowers et al.’s article derived from our work, neither the authors from Nuhfer’s own Idaho State University campus nor the Journal’s editors engaged us in review. The result is an article including several attributions to knowledge surveys and our thoughts/intentions that we disclaim. Readers who compare our ideas about knowledge surveys as presented by Bowers et al. with those published in our words (Nuhfer and Knipp, 2003; Theall et al., 2005; Wirth et al. 2005) should anticipate discrepancies.

REFERENCES


RESPONSE: Re: The Use of a Knowledge Survey as an Indicator of Student Learning in an Introductory Biology Course

Nancy Bowers* and Maureen Brandon†

*Instructional Design and Technology Solutions, Potsdam, NY 13676; and †Idaho State University, Pocatello, ID 83209

We appreciate Nuhfer and Knipp’s comments on our article, and we acknowledge that there are multiple ways to analyze the effectiveness of knowledge surveys as they relate to student learning. Because of the paucity of peer-reviewed research concerning the implementation and interpretation of knowledge surveys, there is no generally accepted method for analysis. The statistical methods we used to address our question of the relationship between confidence levels (e.g., knowledge survey results) and student performance are sound and appropriate. We encourage others, including Nuhfer and Knipp, to publish their findings through the peer-review process, and we look forward to reading those articles.

DOI: 10.1187/cbe.06–07–0173
Address correspondence to: Maureen Brandon (branmaur@isu.edu).

RESPONSE: Re: The Use of a Knowledge Survey as an Indicator of Student Learning in an Introductory Biology Course

Diane Ebert-May and Everett P. Weber

Department of Plant Biology, Michigan State University, East Lansing, MI 48824

What is the correlation between students’ confidence in their level of knowledge and comprehension in a course and their actual performance, as judged by their grades? The research by Bowers, Brandon, and Hill (Bowers et al., 2005) investigated this question using knowledge surveys developed by Nuhfer and Knipp (2003) to determine students’ perceived self-efficacy (belief in one’s capability to carry out an action successfully). A knowledge survey (KS) is a series of content-based questions on topics presented in the course. Students are not asked to answer the questions, but merely to indicate for each question their confidence that they could answer it correctly. Evidence from Bowers et al. did not support the claim by Nuhfer and Knipp that students’ learning can be predicted by perceived self-efficacy levels; rather, their data indicated that the correlation between student confidence and final grades is negligible. In response to this finding, Nuhfer and Knipp challenged the authors’ interpretation of their results (see above letter). We are writing to comment on this controversy and to address a more fundamental question: How useful are knowledge surveys as assessment tools?

DOI: 10.1187/cbe.06–07–0174
Address correspondence to: Diane Ebert-May (ebertmay@msu.edu).

Vol. 5, Winter 2006 315
indeed be a valid link between the KS scores and student understanding reported by Bowers et al.

Reliability and validity are both critical aspects of assessment. Differences in interpretation of assessment results occur and merit productive debate. As Nuhfer and Knipp point out, assessments that define achievement based on lower-level thinking (e.g., knowledge and comprehension) usually have high reliability. If the goal of instruction is for students to demonstrate gains in their knowledge and comprehension of the subject, the KS may be useful. However, science involves more than mastering facts. The KS is not designed to probe students’ confidence about their ability to actively engage in processes of science or higher-level thinking such as analysis or synthesis. As we attempt to assess students’ critical thinking abilities using open-ended problems, determining reliability and validity becomes more difficult (Batzli et al., 2006). Ultimately, we wish to ascertain if assessments reliably measure what we want students to know and be able to do based on the goals and objectives of instruction. Are students becoming more sophisticated in their ability to solve complex problems that do not have single answers, high degrees of completeness, certainty or correctness? Is our instruction providing students guidance and practice in doing so?

Both papers agree that perceived self-efficacy is a key attribute when learning difficult subjects and addressing complex tasks. According to theory, as self-efficacy increases, students are more willing to undertake more complex tasks and think more about complex ideas and problems (Baldwin et al., 1999). Instruction that nurtures critical thinking enables students to gain deeper understanding of content knowledge. Specific guiding questions can facilitate students’ engagement with a problem and draw upon their prior knowledge or identify what they do not know (Ebert-May et al., 2006). Pretests that identify students’ comprehension or misconceptions about a topic can be used to guide instruction.

The peer-reviewed literature about knowledge surveys is sparse, and the relationship of perceived self-efficacy to performance merits further research. However, if critical thinking is the ultimate goal, a KS is unlikely to be a useful assessment. The potential value of knowledge surveys thus revolves around the question of whether “covering” content by the instructor is more important in an undergraduate science course than students “uncovering” content through problem solving and critical thinking. Whether instructors choose to use the KS should depend on their student learning goals. Perhaps instructors need to increase self-confidence in their own ability to promote higher-level thinking by their students.

REFERENCES
Reliability is a fundamental quality of the internal consistence of any measuring instrument. Researchers sometimes use “reliable” and its derivative terms, yet fail to address reliability. Bowers et al. (2005) claimed that the knowledge survey (KS) “does not reliably measure student learning as measured by final grades or exam questions.” They addressed their purpose (p. 311) “to evaluate how closely students’ performance track with their confidence in their knowledge of the course material,” through correlating “plotted pre- and post-KS scores against final grades (p. 314).” This approach assumes that tests/grades of unknown reliability are appropriate standards for judging other measures. Their article offers a case study in drawing conclusions without considering reliability.

The split-halves Spearman-Brown reliability (R) measure (Jacobs and Chase, 1992) derives from the r (r) obtained from individuals’ scores on two halves of a single test. It is a routine method for quantifying reliability and is applicable to both tests and knowledge surveys. When Bowers et al. attributed specific claims to us: “They report that KS results represent changes in students’ learning (p. 311),” and “... are a good representation of student knowledge (p. 316),” they omitted mention of the surprising reliability that characterizes knowledge surveys’ pre- and postcourse measures (Figure 1).

Figure 2 shows a contrasting split-halves reliability for a faculty-made test of unusually high reliability that yields an r of 0.8 and a reliability coefficient of R = 0.9. Good faculty-made tests achieve R > 0.6 (Jacobs and Chase, 1992), but many yield reliability coefficients of <0.3 (Raoul Arreola, personal communication, April 2005; Theall et al., 2005). In correlating the test in Figure 2 with its equivalent KS, the maximum correlation to expect would be about r = 0.8 (the degree to which the less reliable instrument can correlate with itself). We believe r should be still lower, because we developed knowledge surveys to sample cognitive/affective domains that only partially overlap those sampled by tests (Wirth et al., 2005).

Next, consider the reliability of grades derived by averaging several tests. An example appears in Figure 3. Suppose some tests lack significant correlation with other tests, but, as is common practice, we average all the tests anyway to produce grades. The reliability of grades derived from averaging a mix of tests should be lower than the most reliable tests used during the course. If we correlate grades and KS ratings, what, then, are we actually correlating? Without reliability measures, we cannot know.

Short-answer tests longer than 40 items usually have good reliability. Such tests define achievement based largely on knowing, as manifested through short-answer test-taking skills under timed conditions. Tests that address open-ended challenges sample other cognitive/affective domains to define achievement based largely on doing, as manifested through written reports and products generated after discussion, reflection, and revision. Such tests serve as much to promote learning and to mentor students to high-level thinking as to produce grades. High reliability is difficult to achieve using open-ended challenges, and quantifying it requires methods too cumbersome for routine classroom use. Many such tests within Figure 3 contribute to lower overall grade reliability than if achievement were derived solely from short-answer tests. One should know general reliability of one’s tests/grades to understand what comparisons are possible, but it is less important to optimize a correlation between tests and knowledge surveys than it is to mentor students to engage open-ended challenges. Exceptionally effective courses may show near zero reliability in plots such as Figure 3. In such cases, paired comparisons are impossible, but the assessment value of a well-designed KS as a reliable, complementary measure becomes apparent.

Unrealistic expectations for high numerical correlation coefficients between grades and knowledge surveys persist until one understands limits imposed by reliability. Thereafter, understanding permits better interpretations. The positive numerical correlations Bowers et al. (2005) reported as “low” between their postcourse knowledge surveys and grades then seem surprisingly high, given limits imposed by reliability of tests and grades.

We emphasized assessment of learning in classes through use of aggregate data (Nuhfer and Knipp, 2003), which differs from Bowers et al.’s evaluative efforts to predict an individual’s grades from her/his knowledge surveys. At the class level, we were most impressed by Bowers et al.’s Figure 1. It revealed the pattern change from no correlation be-
between grades and the precourse KS to a persistent positive correlation between grades and the postcourse KS. Their tools (tests and the knowledge survey) remained constant through their pre- and postcorrelations, so the profound pattern change seen in every class seems most simply explained by students’ increased understanding of specific content. We suggest correlating the high-reliability pre- and postcourse KS measures as an additional change indicator.

Although Bowers et al.’s article derived from our work, neither the authors from Nuhfer’s own Idaho State University campus nor the Journal’s editors engaged us in review. The result is an article including several attributions to knowledge surveys and our thoughts/intentions that we disclaim. Readers who compare our ideas about knowledge surveys as presented by Bowers et al. with those published in our words (Nuhfer and Knipp, 2003; Theall et al., 2005; Wirth et al. 2005) should anticipate discrepancies.

REFERENCES
RESPONSE: Re: The Use of a Knowledge Survey as an Indicator of Student Learning in an Introductory Biology Course

Nancy Bowers* and Maureen Brandon†

*Instructional Design and Technology Solutions, Potsdam, NY 13676; and †Idaho State University, Pocatello, ID 83209

We appreciate Nuhfer and Knipp’s comments on our article, and we acknowledge that there are multiple ways to analyze the effectiveness of knowledge surveys as they relate to student learning. Because of the paucity of peer-reviewed research concerning the implementation and interpretation of knowledge surveys, there is no generally accepted method for analysis. The statistical methods we used to address our question of the relationship between confidence levels (e.g., knowledge survey results) and student performance are sound and appropriate. We encourage others, including Nuhfer and Knipp, to publish their findings through the peer-review process, and we look forward to reading those articles.

DOI: 10.1187/cbe.06–07–0173
Address correspondence to: Maureen Brandon (branmaur@isu.edu).

RESPONSE: Re: The Use of a Knowledge Survey as an Indicator of Student Learning in an Introductory Biology Course

Diane Ebert-May and Everett P. Weber

Department of Plant Biology, Michigan State University, East Lansing, MI 48824

What is the correlation between students’ confidence in their level of knowledge and comprehension in a course and their actual performance, as judged by their grades? The research by Bowers, Brandon, and Hill (Bowers et al., 2005) investigated this question using knowledge surveys developed by Nuhfer and Knipp (2003) to determine students’ perceived self-efficacy (belief in one’s capability to carry out an action successfully). A knowledge survey (KS) is a series of content-based questions on topics presented in the course. Students are not asked to answer the questions, but merely to indicate for each question their confidence that they could answer it correctly. Evidence from Bowers et al. did not support the claim by Nuhfer and Knipp that students’ learning can be predicted by perceived self-efficacy levels; rather, their data indicated that the correlation between student confidence and final grades is negligible. In response to this finding, Nuhfer and Knipp challenged the authors’ interpretation of their results (see above letter). We are writing to comment on this controversy and to address a more fundamental question: How useful are knowledge surveys as assessment tools?

DOI: 10.1187/cbe.06–07–0174
Address correspondence to: Diane Ebert-May (ebertmay@msu.edu).

Designing assessments that provide substantive feedback about student learning in science presents a difficult challenge to faculty who teach undergraduates. The process of creating and evaluating assessments should include thinking broadly about validity and reliability. Bowers et al. focused on whether KS scores were valid measures of student understanding—validity. The authors concluded that the correlations they found between KS scores and student understanding were too low to validate a link. We concur with Bowers et al. that the statistical methods used in their study were appropriate and that the evidence supported their conclusions. However, Bowers et al. did not address the reliability of their assessments—that is, the reproducibility of the scores that would be obtained if the survey were administered several times to the same students.

Nuhfer and Knipp claimed high internal reliability of their instrument because the scores on related questions within the KS correlated highly with each other. In effect, this implied overall reliability of the instrument. They also suggested that the assessments used by Bowers et al. likely had low reliability, resulting in the observed low correlations between these assessments and the KS. Because of this factor, Nuhfer and Knipp argue in their letter that there may
indeed be a valid link between the KS scores and student understanding reported by Bowers et al.

Reliability and validity are both critical aspects of assessment. Differences in interpretation of assessment results occur and merit productive debate. As Nuhfer and Knipp point out, assessments that define achievement based on lower-level thinking (e.g., knowledge and comprehension) usually have high reliability. If the goal of instruction is for students to demonstrate gains in their knowledge and comprehension of the subject, the KS may be useful. However, science involves more than mastering facts. The KS is not designed to probe students’ confidence about their ability to actively engage in processes of science or higher-level thinking such as analysis or synthesis. As we attempt to assess students’ critical thinking abilities using open-ended problems, determining reliability and validity becomes more difficult (Batzli et al., 2006). Ultimately, we wish to ascertain if assessments reliably measure what we want students to know and be able to do based on the goals and objectives of instruction. Are students becoming more sophisticated in their ability to solve complex problems that do not have single answers, high degrees of completeness, certainty or correctness? Is our instruction providing students guidance and practice in doing so?

Both papers agree that perceived self-efficacy is a key attribute when learning difficult subjects and addressing complex tasks. According to theory, as self-efficacy increases, students are more willing to undertake more complex tasks and think more about complex ideas and problems (Baldwin et al., 1999). Instruction that nurtures critical thinking enables students to gain deeper understanding of content knowledge. Specific guiding questions can facilitate students’ engagement with a problem and draw upon their prior knowledge or identify what they do not know (Ebert-May et al., 2006). Pretests that identify students’ comprehension or misconceptions about a topic can be used to guide instruction.

The peer-reviewed literature about knowledge surveys is sparse, and the relationship of perceived self-efficacy to performance merits further research. However, if critical thinking is the ultimate goal, a KS is unlikely to be a useful assessment. The potential value of knowledge surveys thus revolves around the question of whether “covering” content by the instructor is more important in an undergraduate science course than students “uncovering” content through problem solving and critical thinking. Whether instructors choose to use the KS should depend on their student learning goals. Perhaps instructors need to increase self-confidence in their own ability to promote higher-level thinking by their students.

REFERENCES


Letters to the Editor

How Teaching Matters
Samuel C. Silverstein

Department of Physiology and Cellular Biophysics and Department of Medicine, Columbia University, New York, NY 10032

Harold Wenglinsky’s paper “How Teaching Matters: Bringing the Classroom Back into Discussions of Teacher Quality” (www.ets.org/research/pic) analyzes how the attributes and classroom practices of teachers affect the performance of eighth-grade students on standardized tests in science. By correlating student performance on the 1996 eighth-grade National Assessment of Educational Progress (NAEP) science test with NAEP surveys of the attributes and classroom practices of the same students’ teachers, Wenglinsky identified two teacher attributes and two teacher practices that are highly correlated with superior student achievement.

The two teacher attributes are as follows. 1) Academic preparation of teachers. Students taught by teachers who majored or minored in science or science education in college performed 39% of a grade level above students of teachers who did not have this educational background. 2) Teacher professional development in laboratory skills. Students in classes of teachers whose postcollege professional development included laboratory training performed 44% of a grade level above students whose teachers did not have such training.

The two teacher practices are as follows: 1) Teacher implementation of subject-appropriate and validated, hands-on classroom and laboratory exercises. Students exposed to such exercises weekly were 40% of a grade level ahead of students who participated in such activities monthly or less. 2) Weekly point-in-time short-answer and multiple-choice tests. Students whose teachers used such assessments were 92% of a grade level ahead of those whose teachers used such tests less frequently. According to Wenglinsky, students in classes of teachers who possess both attributes and use both practices were more than a grade level ahead of students in classes of teachers possessing neither attribute and using neither practice. Regrettably, Wenglinsky did not publish these findings in a peer-reviewed journal, so that these attributes and practices are infrequently cited.

Two of the four attributes and practices identified by Wenglinsky (i.e., teacher laboratory skills and implementation of hands-on classroom exercises) are the central focus of Columbia University’s Summer Research Program for Science Teachers (www.scienteacherprogram.org) and of other Science Work Experience Programs for Teachers. Data to be published elsewhere show that teacher participation in Columbia’s program has a very positive impact on their students’ success in passing a New York State Regents exam in science.

If Wenglinsky is correct, science teachers who participate in professional development programs that focus on improving their laboratory skills, and that encourage them to implement more hands-on exercises and weekly point-in-time short-answer and multiple-choice tests in their classes, will be more successful in raising the achievement level of their students than teachers who participate in programs that focus on other aspects of science, education, or both. Confirmation of this postulate could greatly simplify the task of improving middle and high school science education. For this reason alone, it is important to test Wenglinsky’s conclusions rigorously and soon. Although they surely reflect an oversimplification of the measures needed to improve U.S. secondary science education, they are so congruent with the focus of Columbia’s Summer Research Program, and with the very positive impact it has on interest and achievement in science of students of teachers who participate in it, that I felt it important to bring this article and these ideas to the attention of the readership of CBE—Life Sciences Education.

ACKNOWLEDGMENTS
S.C.S. is the John C. Dalton Professor of Physiology and Cellular Biophysics and of Medicine and Director of Columbia University’s Summer Research Program for Science Teachers.
INTRODUCTION

The CBE—Life Sciences Education (CBE-LSE) Approaches to Biology Teaching and Learning feature most often focuses on translating research and scholarship in science education into practical, accessible teaching strategies that can support CBE-LSE readers in their own efforts as science instructors, primarily at the college and university level. In this column, we highlight an issue in science education facing many university and college science departments: hiring faculty who can bring to the department specialized expertise in science education. With increased attention on the scholarship of teaching and on research on teaching and learning approaches unique to individual science disciplines, many science departments find themselves exploring the hiring of a faculty member who is both a scientist by training and a specialist in science education. Although a relatively recent idea in most biology departments, it is increasingly common to find biology, chemistry, geosciences, and physics educators in science departments in colleges across the country. As the visibility of these positions grows, more administrators and faculty in science departments are posing variations on the following questions: How can our science department hire faculty members with expertise in the teaching and learning of our scientific discipline? What roles could this individual play in our department? What background and training would we expect this individual to have? How would the responsibilities for a “science educator” position in our department compare with those of other faculty members? And how do we begin to facilitate a conversation among our science faculty about hiring a specialist in education for our science department?

To begin to address these questions, a collaborative team of tenure-track faculty—all of whom are primarily trained in science and have pursued additional professional development to become education specialists within their discipline—have contributed their collective wisdom on this topic. Among this author group alone, we are aware of more than 18 recent failed departmental searches for these types of faculty positions, indicating the challenges departments face in successfully hiring and retaining Science Faculty with Education Specialties (SFES). Many of these searches remain open because of a lack of candidates matching the advertised qualifications and because of disagreements among science faculty about the nature of the person they want to hire. In addition, there are examples of SFES who have been successfully hired but who leave science departments before tenure because of dissatisfaction with the position. The aim of this article is to share a variety of perspectives on the goals for such appointments and the potential roles of the appointed SFES, providing a document to foster and guide conversations among faculty who are considering hiring a science educator for their science department. The authors of this article represent seven universities and occupy different points along the career trajectory as SFES. The author team includes four assistant professors, one associate professor, and two full professors, of whom five are “biology educators” and two are “chemistry educators.” These authors not only bring their own collective professional experiences to bear, but also those of colleagues from other disciplines and universities who have both informally and formally contributed to discussions on this topic in the context of scientific professional societies and science education communities, such as the American Chemical Society, the American Society for Cell Biology, the Ecological Society of America, the American Physiological Society, and MERLOT (Multimedia Educational Resource for Learning Online and Teaching).

In particular, the authors of this article have been informed by conversations around these questions in two key venues. First, the National Academy of Sciences (NAS) held...
What do SFES look like?

SFES Example 1

George is an assistant professor in the chemistry department. During his doctoral years, George was actively involved in undergraduate teaching and K–12 partnership programs, and he took a variety of science education methods courses and professional development workshops focused on teaching. He now teaches lower-division, nonmajors chemistry courses, an upper-division chemistry majors course, and a methods course for preservice secondary science teachers. George conducts research on the organic photochemistry of carbenes, and in that context he supervises three undergraduate students and collaborates with two of his departmental colleagues.

SFES Example 2

Joel is an assistant professor in the biology department. After receiving his doctoral degree in molecular biology, Joel completed postdoctoral training in science education research. He now teaches lower-division, nonmajors biology courses, two courses that combine methods and subject matter for preservice elementary teachers, and one course that combines methods and subject matter for preservice secondary science teachers. Joel conducts research in biology education by investigating how students interpret static figures and computer animations when learning biology concepts, and he is the first in his department to recruit students into biology education research. He also advises secondary credential students, coordinates the general biology lecture and lab courses, and advises general biology students.

SFES Example 3

Maya is an assistant professor in the physics department. After completing a research master’s degree in physics, Maya completed a doctoral degree in physics education research. She now teaches lower-division, nonmajors and majors physics courses, and a physics content course for preservice teachers in a combined lab and lecture course where she models inquiry and other pedagogical strategies for K–12 classrooms. Maya conducts research in science education, often involving physics content, with a focus on reading and writing in science instruction, and she conducts biophysics research in optical microscopy. She supervises one graduate student and four undergraduate students.

These hypothetical profiles of SFES are examples based on real situations in science departments across a variety of colleges and universities. Typically, the SFES is an assistant professor starting as a science department’s first hire of a tenure-track faculty member with a science education specialty. The typical SFES has formal training in scientific research, sometimes at the master’s level and more often at the doctoral level, as well as specialized preparation in science education through graduate training, postgraduate internships, postdoctoral fellowships, or a second research degree. The lack of a clearly defined career trajectory for SFES results in tremendous variation in the professional backgrounds and training of SFES, especially with respect to the nature of their education and experience in science education. As part of their graduate degrees or during their postdoctoral work, SFES may have completed education courses, engaged in extensive undergraduate or K–12 teaching, obtained a science teaching credential, collaborated on curriculum development, coordinated science education partnerships, and/or conducted science education research. So, with such varied backgrounds and professional responsibilities, how are SFES similar to and different from other science faculty, and how do they fit into science departments?

How are these positions similar to and different from other tenure-track science faculty positions?

As tenure-track faculty in science departments, SFES have the same responsibilities for research, teaching, and service as any other traditional tenure-track faculty member. However, the “education specialty” in SFES positions often leads to two major differences from other faculty. First, the distinctions among the three categories of teaching, research,
and service are consistently less obvious for SFES than for their traditional faculty colleagues; the research, teaching, and service activities of SFES more often overlap. Consider an SFES who develops science content workshops for preservice teachers (undergraduates who aspire to be K–12 teachers) and develops pre- and postassessment surveys to measure gains in their conceptual understanding and identify reasons for those gains. Is this primarily research because the assessment data provide insight into how science is learned and the assessment results can be shared in peer-reviewed forums and used to help obtain external funding? Is this primarily teaching because the SFES taught the workshops and developed new teaching materials? Or, is this primarily service because the broader goal of the project was to increase future teachers’ comfort level with and competency in teaching science? Although this overlap can be viewed positively as a synergistic approach to connecting all aspects of one’s faculty work, these blurred boundaries among the three categories often raise a major challenge to the science faculty’s ability to understand fully the work of the SFES, and thus to be able to evaluate these faculty members for tenure and promotion. In particular, science colleagues often incorrectly assume that this overlap means that SFES will require less time and fewer resources to conduct their teaching, research, and service because they can appear to be doing them all at the same time (see Myths below).

As implied by the three example SFES profiles above, the second major difference between SFES and traditional science faculty is the wide breadth of possible job expectations for SFES hires. Because there is no single model for what an SFES position looks like, administrators and department faculty can easily develop divergent views regarding what they want from such a position. The administration may want the new hire primarily to attend to the broader teaching mission of the university by supporting science teacher preparation programs and being the university liaison striving to improve local K–12 science education. The department may want the new hire primarily to institute best teaching practices by coordinating and training instructors in the introductory science courses, whereas the new hire may primarily want to conduct, report, and peer review science education research or basic science research, all the while building networks and establishing credibility with peers. Also, in some cases, SFES may hold a joint appointment, such as a 50/50 split faculty position between a science department and an education department. These joint appointments across colleges or departments lead to an even wider set of possibilities and consequent expectations.

**On Research Expectations**

SFES research interests and activities are as varied as those of their traditional tenure-track counterparts, but the variety of activities that may be considered scholarly activities for SFES can be extreme. Does SFES scholarship mean core science research, such as Joel’s work in photochemistry of carbenes? Does it mean education research strongly connected to the discipline, such as George’s studies of visual approaches in biology instruction? Does it mean research more broadly related to science education, such as Maya’s investigation of reading and writing in science instruction? Could scholarly activities be interpreted as assessing and improving teaching practice at the department level, or supporting K–12 teachers in their training and professional development? Grappling with what is meant by SFES scholarly work occurs within the context of a movement toward a broader definition of scholarship in the scientific disciplines, one that also embraces discoveries about teaching and the generation of new knowledge about educational approaches within the discipline (Boyer, 1990). Finding and implementing the best instructional approach fits with other definitions of scholarship, such as the scholarship of integration, the scholarship of application, and the scholarship of teaching (Boyer, 1990). This practical work draws upon the knowledge gained through science education research. In the current funding environment, significantly more support is available for incorporating and disseminating best practices in undergraduate and K–12 science education than for pure discovery projects. In addition, basic science education research is meant purely to create new knowledge, much like basic science research, and is unlikely to take the form of any immediate application, such as a solution to a specific problem in teaching practice. This pure form of science education research often goes unacknowledged by traditional faculty who may assume that an SFES would engage exclusively in research specifically applicable to teaching situations in their own department.

Given the breadth of research that SFES may wish to pursue, three types of SFES research expectations deserve a closer look: basic science research, science education research, and a combination of these two fields. SFES engaging in basic science research, such as George studying carbenes and Maya investigating optical microscopy, may have more common professional ground with their department colleagues in terms of their research questions and methodologies. They will generally have the same external validation expectations and venues in terms of publications and presentations, and the department will likely have a better ability to appreciate and evaluate their research activities. However, maintaining equitable distribution of responsibilities among research, teaching, and service is a challenge here, because the expertise that SFES bring to the department may be less about their research area, and more about their science education background. As mentioned previously, the nature of the education specialty of SFES can easily lead to greater teaching and service commitments than those of other entering faculty, thus overextending the SFES at a critical point in their budding research career.

SFES engaging in science education research, such as Joel and Maya, fit the NAS definition, that SFES positions are science education research positions intended to build knowledge about the teaching and learning of a particular scientific discipline (NAS, 2006). External validation expectations for this group are similar to their traditional tenure-track peers. These expectations include: peer-reviewed dissemination (publication and meeting participation), activity within a science education community, writing competitive science education research grants, and garnering external funding for science education projects. Although SFES build knowledge about the teaching and learning of their particular scientific discipline and in some cases directly (and positively) impact students and faculty, there are several cultural challenges for SFES doing only science education research.
research in a science department. Many natural scientists are unfamiliar or uncomfortable with social science research, and even suspicious about these endeavors, particularly nonquantitative research methodologies such as interview techniques, case studies, or grounded theory research. Further, they may possess the misconception that science education research is somehow easier, quicker, and cheaper than basic science research. Those SFES engaged in purely science education research face the challenge of building their research program in a context where they have few to no department colleagues with whom they can discuss their research and potentially have colleagues who do not even acknowledge their work as research.

Finally, engaging in both science and science education research can give the individual SFES the greatest latitude in pursuing their research interests. This type of position may be particularly attractive to candidates making the transition from more traditional scientific research paths to science education research careers. However, a significant drawback is the fact that developing one research program that is outstanding and rigorous is challenging enough, let alone attempting to simultaneously develop two different research programs—one in basic science and one in science education. In this situation, SFES, who are usually junior faculty, must grapple with the daunting task of developing two lines of research inquiry, each with its own methodologies, literatures, professional meetings, circles of colleagues, and cultures of external validation that are likely to be nonoverlapping. Finally, when there are possibilities of SFES doing research in two different areas, there is also greater risk of divergent expectations between the department and the SFES hire.

**On Teaching Expectations**

In some cases, SFES teaching assignments may be similar to those of traditional tenure-track peers, in that they can span the range of a department’s offerings from the introductory to the advanced level. However, they can sometimes be very different. SFES commonly teach or supervise large-enrollment courses, such as lower-division introductory courses for science majors and introductory service courses for non-science majors, which primarily may be taught by lecturers at large-enrollment universities. In some instances, SFES teach upper-division courses in the department in the sub-discipline in which they were originally trained (e.g., microbiology). In contrast to their traditional peers, SFES are also frequently asked to teach discipline-specific courses for preservice teachers. These may include broad content survey classes for preservice elementary school teachers and science teaching methods courses for preservice secondary school teachers. Further, some SFES teach courses on science education research methods and discipline-specific teaching methods for upper-division or master’s-level students in the department, including teaching assistant training courses. Thus, SFES often teach more large-enrollment courses and a greater variety of courses than their peers. In addition, SFES are frequently expected—implicitly or explicitly—to engage in a greater degree of course, curriculum, and program development and may be heavily relied on by other faculty because of their science education expertise. Finally, it may also be expected that an SFES will somehow transform a department’s overall educational approach or at the least act as a catalyst to affect educational reform within a department. These combined teaching expectations, although all appropriately aligned with the expertise that SFES bring to a department, are in total a heavy load for any junior faculty member developing his or her academic career.

**On Service Expectations**

Because SFES are usually the only—or at best one of a few—science education specialists in their department or college, they are often tapped to serve on any committee that has even a slight relationship to education, science education, and/or assessment at the department, college, or university level. At larger institutions, many SFES also may have course coordinator roles for introductory majors courses, including the responsibilities of hiring, training, and supporting teaching assistants and adjunct instructors. They may often play a greater role in departmental assessment activities and program review processes and are looked to for immediate leadership on departmental curriculum committees. In addition, SFES may be drafted to take on outreach and partnership responsibilities with local K–12 teachers and schools. They are commonly asked to play major roles in K–12 teacher preparation, recruitment, and retention. Further, they may be expected to contribute to preparing state-mandated teacher preparation accreditation materials, which often involves working with many science faculty to review the department’s academic program and its alignment with newly minted teacher preparation standards, a major undertaking. In this same vein, SFES frequently serve as community science education resources, for example, supporting local science fairs or regional and national precollege science competitions. Some SFES have extensive ambassadorial responsibilities, including running campus tours of science facilities or hosting visits from science clubs and student organizations. Again, there is a serious potential problem of overloading a junior faculty member at the start of his/her career.

**WHY CREATE THESE POSITIONS WITHIN SCIENCE DEPARTMENTS?**

Although a common initial reason for hiring SFES in science departments is to benefit the department’s own curriculum (particularly improving lower-division courses), undergraduate teaching and curriculum needs are only one of several rationales for hiring SFES. Because the rationale for hiring an SFES will begin to dictate the expectations the department sets for this individual, clarity on why an SFES is being hired into a science department is critical. The four most prevalent rationales for hiring SFES in a science department are summarized below.

1. Education research on the teaching and learning of a discipline is an important area of study within that discipline, for example, research in biology education is a sub-discipline of biology. Such research builds discipline-specific knowledge about effective teaching and learning of that discipline. The mental pathways and models underlying deep conceptual understanding of biology are different from those underlying understanding of geol-
ogy, which are different from those underlying understanding of physics, etc. Misconceptions that challenge students in a particular discipline are best studied by faculty with deep understanding of the content and methods of the discipline. Faculty based in colleges of education often focus less on research that is informed by advanced knowledge of a science subspecialty area; their research tends to focus, instead, on cross-cutting issues such as equity, the nature of science, and assessment, that is, topics that are not specific to biology, chemistry, etc.

2. Close contact with SFES can enhance collaboration in the department among faculty contributing to advancement of undergraduate science education and can improve funding prospects. The importance of discipline-based knowledge in science education has led the National Science Foundation (NSF) to fund grants that broadly impact teaching and learning (see Table 1), and so opportunities exist for all faculty to share and apply their expertise in curriculum development. To merit NSF funding, all proposals must meet two basic criteria: intellectual merit and broader impact. SFES typically have training in addressing both with special expertise in broader impact. Of the 40,000 proposals submitted to NSF each year for research, education, and training projects, approximately 11,000 are funded. For every 100 proposals received, ~80 have good ideas (intellectual merit), but only ~25 have good ideas, are well written, and have a good realization of broader impacts. To show broader impacts, NSF-funded research should "advance discovery and understanding while promoting teaching, train-

### Table 1. Selected funding sources for science education research

<table>
<thead>
<tr>
<th>Agency</th>
<th>RFP grant title</th>
<th>Grant duration</th>
<th>Funding level</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSF</td>
<td>Research and Evaluation on Education in Science and Engineering (REESE)*</td>
<td>Synthesis research: 1–3 yr</td>
<td>Up to $200K</td>
</tr>
<tr>
<td>NSF</td>
<td>Nanotechnology Undergraduate Education Course, Curriculum, and Laboratory Improvement (CCLI)</td>
<td>Empirical research: 3–5 yr</td>
<td>Up to $1 million</td>
</tr>
<tr>
<td>NSF</td>
<td>Informal Science Education*</td>
<td>2 yr</td>
<td>Up to $200K</td>
</tr>
<tr>
<td>NSF</td>
<td>Faculty Early Career Development Program (CAREER)*</td>
<td>Phase 1: 1–3 yr</td>
<td>$150–200K</td>
</tr>
<tr>
<td>NSF</td>
<td>Research on Gender in Science and Engineering</td>
<td>Phase 2: 2–4 yr</td>
<td>Up to $500K</td>
</tr>
<tr>
<td>NSF</td>
<td>Discovery Research K–12 (Applied Research, Development of Resources and Tools, and Capacity Building)*</td>
<td>Phase 3: 3–5 yr</td>
<td>Up to $2 million</td>
</tr>
<tr>
<td>NSF</td>
<td>Science Education Partnership Awards</td>
<td>Project grants: 1–5 yr</td>
<td>$100K to $3M</td>
</tr>
<tr>
<td>NSF</td>
<td></td>
<td>Planning grants: 1–2 yr</td>
<td>Up to $75K</td>
</tr>
<tr>
<td>NSF</td>
<td></td>
<td>Workshop grants: 1–2 yr</td>
<td>$50–200K</td>
</tr>
<tr>
<td>NSF</td>
<td></td>
<td>5 yr</td>
<td>$400K+</td>
</tr>
<tr>
<td>NSF</td>
<td></td>
<td>Research awards: 1–3 yr</td>
<td>$500K/yr</td>
</tr>
<tr>
<td>NSF</td>
<td></td>
<td>Dissemination awards: 1–3 yr</td>
<td>$200K/yr</td>
</tr>
<tr>
<td>NSF</td>
<td></td>
<td>Extension services: 5 yr</td>
<td>$500K/yr</td>
</tr>
<tr>
<td>NSF</td>
<td></td>
<td>Conference: 2 yr</td>
<td>Up to $100K</td>
</tr>
<tr>
<td>NSF</td>
<td></td>
<td>Exploratory study: 3 yr</td>
<td>Up to $300K</td>
</tr>
<tr>
<td>NSF</td>
<td></td>
<td>Full-scale study: 5 yr</td>
<td>$1.5–3 million</td>
</tr>
<tr>
<td>NIH: NCRR (Natl. Center for Research Resources)</td>
<td>Science Education Partnership Awards</td>
<td>Phase 1: 1 yr</td>
<td>Direct costs up to $250K/yr</td>
</tr>
<tr>
<td>Knowles Foundation</td>
<td>Knowles Scholar Awards</td>
<td>2 yr</td>
<td>$110K</td>
</tr>
<tr>
<td>Knowles Foundation</td>
<td>Math and Science Education Research Awards</td>
<td>Up to 4 yr</td>
<td>$150K–$750K</td>
</tr>
<tr>
<td>Department of Education</td>
<td>EPSCoR Comprehensive Program</td>
<td>3 yr</td>
<td>$150K–$600K</td>
</tr>
<tr>
<td>Department of Education</td>
<td>The Mathematics and Science Partnership (MSP) Program</td>
<td>The state agency for higher education provides competitive awards made to partnerships to improve teacher knowledge in mathematics and science. Determined by state agency (b)</td>
<td></td>
</tr>
<tr>
<td>Department of Education</td>
<td>Academic Improvement and Teacher Quality Programs</td>
<td>The state agency for higher education provides competitive grants to partnerships comprising, at a minimum, schools of education and arts and sciences along with one or more high-need local educational agencies. Determined by state agency (b)</td>
<td></td>
</tr>
</tbody>
</table>

* For fiscal year 2007, the NSF Divisions that administer these programs (Research, Evaluation, and Communication [REC] and Elementary, Secondary, and informal Education [ESIE]) are scheduled to be combined into a single division, Research on Learning in Formal and Informal Settings (DRL).

\(b\) To identify your state MSP coordinator, see [http://www.ed.gov/programs/mathsci/faq.html](http://www.ed.gov/programs/mathsci/faq.html), and for information on Academic Improvement and Teacher Quality Programs (AITQ), see [http://www.ed.gov/about/offices/list/oe/se/aitq/aboutus.html](http://www.ed.gov/about/offices/list/oe/se/aitq/aboutus.html)
ing, and learning” (NSF, 2004, p. 39). Projects funded by the Division of Undergraduate Education devote part of their budget toward evaluation to document any behavioral changes in faculty or students that have led to more successful learning of science material. Proposals from a department with faculty (e.g., SFES faculty) who are able to conduct this evaluative research on outcomes from activities associated with proposals for equipment or basic research grant proposals are more likely to be viewed as meritorious, especially if the documented behavioral changes from prior projects have proved to be sustainable. SFES in science departments may additionally use their experience with social science research methods to identify ways to broaden the participation of underrepresented groups (e.g., gender, ethnicity, disability, geographic, etc.) and enhance the infrastructure for research and education through networks and partnerships.

3. Faculty in science departments are ultimately the ones responsible for the science knowledge of K–12 teachers, yet science departments are often unhappy with the quality of K–12 science teaching. University faculty set the standards for the entire educational system, and if science faculty aspire to improve K–12 science education, they must identify and adapt instructional strategies that engage and support the diversity of students in college classrooms. In other words, science faculty can specifically support science education by how they teach future science teachers. SFES can foster connections between science faculty and K–12 science education through training components for preservice teachers, including methods courses and student advising.

4. SFES bring expertise to a science department’s program review, curricular reform, and accreditation processes. The special expertise of science faculty who understand how people learn science and how to use formative assessment and rubrics to help students meet goals, how to clearly define objectives, and who are aware of challenges faced by diverse learners is essential if we value equitable access to the knowledge of science. This is a huge topic for any campus faced with program review or regional accreditation, because addressing the needs of diverse populations of students and institutionalizing program assessments have become important foci of the accreditation process.

COMMON MYTHS ABOUT SFES POSITIONS

Up to this point, we have considered what SFES look like and why departments should hire such faculty. Efforts by many science departments to bring science education specialists into the science disciplines have been successful. However, for every successful integration of an SFES into a department, there appears to be an unsuccessful situation, in which either no SFES was hired from a search or an SFES was hired but left the positions because of various difficulties. The authors of this article have observed that a common set of incorrect assumptions about the nature of SFES positions may seriously hinder successful hiring and integration of SFES into science departments. Addressing these assumptions is critical if scientists are to embrace science education within their discipline. We highlight six myths we find to be pervasive that can derail the hiring and support of SFES. Our descriptions are modeled after the Bower report (Bower, 2005), written to highlight common myths about K–12 education based on an 11-year collaborative partnership between Caltech and the Pasadena Unified School District.

Myth 1

The new science faculty we hire who specialize in science education will develop new courses and add innovations to existing courses that will solve the problems incurred in teaching nonscience majors who are deficient in their understanding of basic science.

Developing curriculum is a huge and time-intensive endeavor. James Bower, a pioneer SFES, admits the following (Bower, 2005):

The most important personal consequence of my involvement with science education reform has been a growing awareness of how poorly I have taught my own students (cf.) . . . After 10 years of involvement with precollege science, I have become profoundly aware of the negative effect the poor teaching of science in colleges and universities has on the rest of the educational system.

If nonscience majors are deficient in their understanding of science, one course with science curriculum designed by one person has little potential to alter the cycle of ignorance. Science education reform requires a concerted effort by groups of faculty who work together over time. For example, integration of lab activities with lecture components of courses that serve large groups of nonscience majors will happen only if large and diverse groups of faculty collaborate to improve the courses. Some experienced science faculty who are thoroughly familiar with the needs and motivations of local students may not be open to innovations meant to involve students with content in ways that are meaningful and relevant. Not all SFES are interested in or qualified to conduct faculty professional development to assist even receptive colleagues to rethink the way they teach, let alone more resistant ones. However, although it may be a myth that SFES can make all the changes needed to help remediate all students’ deficiencies, they can play a vital role by providing their insights to departmental curriculum reform efforts.

Myth 2

New faculty who specialize in science education can replace senior faculty who have dedicated their careers to teaching and furthering the causes of science education.

Senior faculty who specialize in science education work under very different conditions than junior faculty who are hired as SFES. First, from their experience, such senior faculty know local students’ motivations, cultural context, and limitations. Second, senior faculty are typically tenured and supported by a network of colleagues in their discipline who provide access to knowledge and instructional resources. After spending years working on instructional problems, they have found resources about topics they are expected to teach that may not have been dealt with in their own re-
search or educational training. They have most likely introduced innovations that match the needs of the particular students in their courses. Junior faculty who specialize in science education have the same need as other faculty to conduct and peer review research in order to build faculty support networks that they can access when they encounter problems with diverse new teaching experiences. Given the chance to build academic networks, careers can develop over time leading to a position where the SFES can continue to learn and contribute. It is unreasonable to expect knowledge of science education research to cause some faculty to start out with the local expertise of senior faculty who chose to specialize in science education and have gained years of teaching experience.

Myth 3
By designing and teaching a few courses that align goals with appropriate activities and assessments for students who will complete teacher certification programs, faculty who specialize in science education can become responsible for the science training of future K–12 science teachers.

The science knowledge, skills, habits of mind, and experiences expected for outstanding K–12 science teachers are not trivial. Every science faculty member has a stake in producing good K–12 teachers, and it is unreasonable to think that this goal will be met by modeling appropriate teaching strategies in just a small subset of undergraduate courses. SFES may be highly qualified to help other faculty set goals for the science elements of a teacher credential program. They can assist in matching goals to appropriate instructional activities with both formative and summative assessments to make sure teacher accreditation standards are met. However, it should be up to each individual faculty member to define specific instructional activities and objectives to help teacher candidates meet science education goals, including familiarity with conceptual understandings, scientific investigation, lab skills, problem solving, communication strategies, teamwork building, practical reasoning, and positive attitudes toward science. Although it is unreasonable to have one faculty member take charge of all the instructional components appropriate for teacher credential candidates, a science department will be enriched by having a colleague who understands teacher credential requirements and who can model appropriate pedagogy for future K–12 teachers in science content courses.

Myth 4
Faculty who specialize in science education can teach and provide service to the department without the need for resources (time and space) to do research because their research is conducted as they teach their classes.

Rigorous education research requires a representative sample of students that is not often provided by the students in one’s own classes. When research is conducted in a “convenience sample” of students in one’s own course, then extensive qualitative work is required to document the student demographics so that others will know to what degree results are of interest and can be generalized to other situations with different student populations. Sample size is another problem. The academic year can be a severe limitation when an entire year must go by before a study can be replicated. Thus, it is a myth that SFES need only limited time and space for research. Any faculty who are denied the time and resources to conduct research, including SFES, will be limited in the research they can do and in their development as a scholar, which will in turn impede their effectiveness in teaching and serving the department. As such, resources for SFES research can be seen as an investment that leads to local educational expertise within the faculty.

Myth 5
The problem with science education in general is that social science research methods and science education theory lack the rigor of basic science.

Few basic science faculty have much experience with qualitative data and the social science methods used to answer important questions in science education, and so they are not highly qualified to judge the academic rigor of science education studies. When investigating science education, random assignment of students to treatment groups is often not feasible and can even be unethical. In some cases, focus group discussions or solicitations of student input through surveys or interviews can be much more rigorous and effective than conducting experimental education research. Methods such as triangulation, bricolage, and other research techniques provide rigor for the empirical basis of science education theory that is supported by vast bodies of evidence. Examples of the use of these research methods to solve important instructional problems are apparent from a Google Scholar search of the terms. Among education researchers, the rigor of education research methods has recently been a topic of some debate. (For a summary of the issues, refer to Pellegrino and Goldman, 2002; Shulman et al., 2006.) To help faculty navigate the complexity and assess the controversy surrounding results from social science research methods, reports such as “How People Learn” from the National Academies have been cautiously written using standards of empirical evidence meant to report only consensus ideas about how people learn (National Research Council, 2000). Having a faculty member in the science department who keeps up with advances in science education research can help maintain a departmental focus on proven methods while protecting the department from failed innovations resulting from unproven fads or political directives.

Myth 6
An SFES will be the best teacher in a science department and will have nothing to learn from the teaching experiences of senior science faculty.

The majority of SFES are junior faculty who have just emerged from some specialized educational training after a traditional graduate research experience in their home scientific discipline. Although these individuals may have taught for several years, they cannot bring to the department the decades of classroom experience that many senior science faculty (albeit with no special training in science education) have accumulated. That said, a common myth that is pervasive and divisive is the assumption that SFES can and will immediately outperform experienced faculty and
emerge as the best teacher in the department. Not only is this unlikely, it sets up a false dynamic that does not acknowledge the science teaching strengths of the basic science faculty, and at the same time it puts unreasonable expectations on these young faculty members who are at the beginning of their pedagogical careers. Experienced science faculty should not feel threatened by these new “teaching experts.” In fact, experienced science faculty and new SFES faculty have complementary expertise that, when shared, will be mutually beneficial and open up conversations about teaching and learning that may not have previously occurred within a department.

Although we share concerns about these six myths, it should be noted that behind each myth is a kernel of truth about what a colleague who specializes in science education can actually do for a science department. A department that hires such an individual can expect that individual to help the department to engage and serve the needs of diverse students and to examine and improve instruction to better meet department goals. SFES can participate in faculty collaborations to improve courses and programs for K–12 science teachers to break the cycle of ignorance. As more teachers understand how people learn science, increasing numbers of children will be empowered by the knowledge of that discipline. SFES can help build strong collaborations among faculty who continually innovate with instructional methods to make the science discipline more deeply meaningful to students. They just can’t do it all at once!

WHAT CONVERSATIONS DOES A DEPARTMENT SEEKING TO HIRE A SCIENCE FACULTY MEMBER WITH AN EDUCATION SPECIALTY NEED TO HAVE?

As emphasized in the introduction, there are few to no resources available to aid science departments and colleges of science aspiring to hire a scientist who specializes in science education (NAS, 2006). With more departments opening searches for these SFES, there is an emerging need for guidelines for departments embarking on discussions of how to define and search for someone to fill such a position in their discipline. Given the dearth of guidelines, we have developed a tool to aid departments in beginning these conversations, a Departmental Guide for Discussions on Hiring SFES, which is available in the Supplementary Material. This guide poses questions that we believe can aid science departments in engaging in conversations that: 1) establish the goal(s) of hiring one or more SFES; 2) articulate both the science and science education background and training desired in an SFES; 3) delineate expectations for research, teaching, and service for an SFES; and 4) anticipate what implications the hiring of SFES has for the department’s future curriculum, hiring plan, and academic programs and degrees. These points for departmental conversations, described in more detail below, are not intended to be complete, but rather to be the starting points for discussion.

Establish the Goal(s) of Hiring One or More SFES

In evaluating potential new colleagues in traditional areas of research in a science discipline, science faculty often share common visions about expectations in research, teaching, and service. However, faculty have much less experience in hiring SFES who are scientists by training but who also specialize in science education. Search committees, and departments as a whole, need to articulate their goals in hiring an SFES and clarify the responsibilities of SFES hires in their departments before the hiring process and to be prepared to communicate these expectations explicitly to SFES candidates. Brainstorm early about goals for hiring an SFES and define department needs well before writing any job description. Being able to articulate goals can lead a faculty to their ideal job description and will allow hiring committees to both explicate the hiring criteria and also begin to discuss the evaluation process for what is, in most cases, a novel position. The process of developing a consensus on goals should also lead to realistic expectations.

Once the goals of hiring an SFES are somewhat defined, take the time to reach consensus and review these goals with all departmental faculty and administrators. A process of collaboration and consensus building about the nature and goals of an SFES position among administrators and faculty will highlight the valuable contributions such a colleague can bring to a science department and dispel incorrect assumptions, such as the six myths described above. In addition, widespread agreement on the goals of hiring an SFES from the outset will help departments to avoid the unfortunate situations in which 1) the search committee has an idea of the “ideal candidate” that doesn’t match the ideas of the rest of the interviewing faculty or 2) where the “ideal candidate” is hired, based on the search committee’s criteria, but other departmental colleagues have different criteria in mind, causing discord later in the tenure and promotion process.

Articulate Both the Science and Science Education Background and Training Desired in an SFES

Depending on the goals of the department for an SFES position, the background desired may differ substantially. In crafting an SFES position and job description, indicate the types of science and science education training that are most desired by the department—whether it be master’s-level or Ph.D.-level basic research training, extensive undergraduate teaching experience, formal education research training, or experience interacting with the K–12 school system. Also, depending on the desired SFES qualifications, departments may be well served by inviting external stakeholders into the discussion and perhaps asking them to serve as members of the search committee. Potential stakeholders to include in the dialogue include SFES already at the university in other science departments, faculty in the College of Education who are potential colleagues of this new SFES, and leaders from the local K–12 school systems.

Delineate Expectations for Research, Teaching, and Service for an SFES

As stated earlier, a wide range of expectations exist for SFES, so departments must carefully consider their goals and define expectations that a new faculty member can reasonably achieve. The questions in the Supplementary Material aim to push departments to do so and to explicitly discuss how these SFES expectations will compare with those for
Anticipate the Implications of the Hiring of an SFES for the Department’s Future

In crafting SFES positions, science departments may find their expectations far exceed the capacity of one SFES position. A plan for hiring multiple SFES over a period of time may aid the department in prioritizing its goals for the initial and subsequent SFES hires. These departmental discussions would likely move toward how the hiring of one SFES may affect many aspects of the department in the future. Future implications include defining SFES tenure and promotion criteria and establishing undergraduate and graduate curricula and programs in science education, including science education research.

In using the Departmental Guide for Discussions on Hiring SFES (Supplementary Material), we hope that science faculty will be able to openly examine the many options for defining these positions and explore how SFES can enhance the connections between the scholarship of science teaching and learning and the science disciplines on their campus.

MOVING FORWARD

In closing, we encourage you, the reader, to consider how the information contained in this article could be shared with your departmental colleagues and to reflect on how you might support the integration of an SFES in your department.

- Educate yourself and prepare to validate the expertise of your department’s SFES, especially if you notice that other faculty in your department make incorrect assumptions about their science education expertise or the ways in which their time should be spent.

- Help your department distribute service equitably (or in line with the original, stated expectations) so that faculty members hired as SFES are not unexpectedly asked to assume more service requirements than faculty who do basic research in your discipline.

- Advocate for science education research time and space so that science faculty who specialize in education can create new knowledge and subject their work to peer review as they interact with and establish their credibility among peers.

- Find opportunities to collaborate on courses directed by faculty who specialize in science education. In the process you may learn about innovative science teaching approaches, and you can contribute your disciplinary expertise to improve the courses.

- Encourage all faculty in your department to participate in science education reform at all levels by working together to define goals, identify activities appropriate for specific objectives, and create assessments to measure the degree to which your students master the elements of knowing and doing science, including conceptual understanding, scientific investigation and lab skills, problem solving, communication, teamwork, safety, practical reasoning, and positive attitudes toward science.

Finally, we hope that you are able to use the ideas presented in this article to encourage a conversation in your own department about hiring, retaining, and supporting SFES. If your department finds that more support, guidance, and leadership is needed from higher levels, such as the university administration or the professional societies in your discipline, use your voice to help clarify the opportunities and needs we have illustrated here. The increased hiring of SFES we have described holds great promise to enrich science departments and contribute to improved science education in a variety of ways, but only if departments are successful in crafting appropriate positions and hiring, supporting, and retaining these science faculty with education specialties.

ACKNOWLEDGMENTS

The authors thank the California State University Chancellor’s Office for funding the 2006 CSU Science Education Colloquium and acknowledge the contributions and insights of all those who attended that event. N.P., on leave from CSU, Fullerton, is supported by the NSF Independent Research/Development Program while serving at the Division of Undergraduate Education (EHR/DUE) of the NSF. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES


The human brain contains an estimated 100 billion neurons, and browsing the Web, one might be led to believe that there’s a Web site for every one of those cells. It’s no surprise that there are lots of Web sites concerning the nervous system. After all, the human brain is toward the top of nearly everyone’s list of favorite organs and of intrinsic interest to scientists and nonscientists alike. In this review I will focus primarily on three excellent Web sites that represent three very different approaches to Web design and educational media. At the end of this review, I list a number of noteworthy neuroscience Web sites but without extensive remarks.

THE ILLUSION OF SCIENCE

To begin, as a reminder of the wonder of the human nervous system, let’s consider some optical illusions. A commercial site, Sandlot Science (http://www.sandlotscience.com/; Figure 1), has one of the best collections of optical illusions on the Web. The illusions are optimized for viewing on the Web and include interactive Flash-based features, as opposed to simply presenting flat-art illustrations on a Web page. I recommend taking one of their tours. Sandlot even provides an estimate of how many minutes each exploration will take. One of the most perplexing illusions is called the “Mysterious Mind Tap,” number 4 on Tour 1. I look forward to seeing readers’ explanations of how this amazing trick works. Like most optical illusion Web sites, including those posted by academics, Sandlot is frustratingly short on rigorous explanations of how the illusions trick our visual systems. In part this is a good lesson on how far we have to go in understanding even such a well-studied nervous component as the visual system.

Another commercial Web site, Grand Illusions (http://www.grand-illusions.com/), sells some interesting toys and has first-rate illusions freely available on the Web site. They also have some nice video demonstrations and explanations. Two of my favorites are the Dragon Illusion and Dr. Angry and Mr. Calm. The Dragon Illusion (http://www.grand-illusions.com/opticalillusions/dragon_illusion/) is illustrated with a very effective online video. In case you are wary of camera tricks, you can also print out the illusion, paste it together, and check it out for yourself; it really works. I have made dragons of different sizes, and the illusion is scalable. I’d like to see a huge version in a town square or a mall. Dr. Angry and Mr. Calm (http://www.grand-illusions.com/opticalillusions/angry_and_calm/; Figure 2) is a very effective illusion that works well online and might also work in a classroom as long as students are free to move about the room. Are my eyes deceived, or should this illusion be renamed Dr. Angry and Ms. Calm? Hopefully, playing with optical illusions will get students thinking about how the visual system and other brain centers process information. Students might even be wondering whether it’s the eyes or higher brain centers that are being tricked. The University of Wisconsin has a Web site (http://www.physiology.wisc.edu/yin/public/) that contains amazing videos of original Hubel and Wiesel experiments on cat receptive fields. Seeing how receptive fields work can lead one to think about how such mechanisms could be fooled by various visual features in the environment.

Before we leave sensory systems, I’d like to recommend at least one comprehensive site on the visual system, put together by three faculty members at the University of Utah John Moran Eye Center. WEBVISION (http://webvision.med.utah.edu/) is not particularly interactive, but it is very well organized and features some beautiful illustrations and photographs. Keeping in mind that we have other sensory systems, have a look at Dangerous Decibels (http://www.dangerousdecibels.org/virtualexhibit.cfm#) built in partnership by Oregon Museum of Science and Industry and Oregon Health and Science University. The site aims to encourage young people in particular to take care of their hearing, and it includes some interesting simulations of what things sound like under various conditions of hearing impairment. It had me thinking twice about the volume on my iPod.

THE BRAIN FROM TOP TO BOTTOM

The great challenge, and advantage, in teaching about the nervous system is the huge diversity of topics and levels of analysis in neuroscience. Neuroscience researchers are supreme reductionists, down to their favorite molecules. Neuroscientists are also interested in emergent properties of mind and even social systems. Their toolkit ranges from DNA sequencer to computer-assisted analysis of behavior. Perhaps no Web site more overtly embraces the full range of

DOI: 10.1187/cbe.06–02–0147
Address correspondence to: Dennis Liu (dliu@hhmi.org).
neuroscience analysis than The Brain From Top to Bottom (http://www.thebrain.mcgill.ca/flash/index_d.html; Figure 3). In fact the very design of the Web site attempts to epitomize the various levels of neuroscience analysis.

The Brain From Top to Bottom was developed, researched, written, and “copylefted” by Bruno Dubuc and hosted by McGill University (sponsored by the Canadian Institutes for Health Research and the Canadian Institute of Neurosciences, Mental Health, and Addiction). Copyleft means that the content is free for reuse and even modification; however, other than some quirky creativity, the site is focused not on the politics of intellectual property, but on neuroscience. This handsome, well-designed site is worth looking at first and foremost for its ambitious attempt to promulgate a design that organizes material by topic: Functional Anatomy, Memory, Pleasure and Pain, Emotions, Brain Evolution, Movement, The Senses, and Mental Disorders (Development, Language, Sleep and Dreams, and Consciousness are “under development”). Via parallel threads by level of biological organization, the site offers Social, Psychological, Neurological, Cellular, and Molecular content. Third, the material is presented according to color-coded options of Beginner (orange), Intermediate (teal), and Advanced (red) levels of explanation. As if that weren’t enough, the Web site is available in French and English. Viva la synapse!

How does this complicated-sounding scheme work? For example, under the topic of “Memory”—advanced explanation—social level, we find a discussion of Dawkins’ memes concept and the transmission of culture. Shifting to beginner explanation, oral tradition and writing are discussed. Moving to the psychological level—beginner explanation, short-and long-term memory are covered, and more detailed flow-charts of memory are presented along with more technical descriptions at the advanced level. At the molecular level, long-term potentiation is covered for beginners and gluta-mate and synapses are mentioned. Intermediate explanation adds information on receptor subtypes, and advanced explana-tion gives details of second-messenger cascades and gene activation.

The core Web site content is bolstered by five categories of Supplementary Information: Experiment, History, Linked, Tool, and Research. I found the supplemental categories a bit confusing because the Experiment category did not necessarily describe specific experiments, and the Tool modules didn’t discuss methods or techniques. That said, the short Tool module, for example, on the relationship between ontogeny and phylogeny and the “recapitulation” argument is one of the better short explanations of Haeckel’s natural law. The site also features provocatively titled Guided Tours: “Why do I say so many stupid things when I get drunk?” “Why do I make myself sick, instead of just strangling my boss?” “Why can I remember exactly what I was doing the morning of September 11, 2001?” “What’s the connection between the Big Bang and a brilliant chess move?” These questions will draw you and your students in for self-directed learning.

The emphasis of the Web site is good design, clear graphics, and engaging explanation. Because of this emphasis, I think this site is particularly worthy of study by faculty and other independent, small-scale, developers of learning media. From a technology and resources viewpoint, anyone could make a site like this. All the energy and resources have
gone into the design, functionality, esthetic, and content. Given the interesting design of this Web site, I look forward to continued addition of neuroscience topics and subtopics and perhaps the incorporation of some interactive media.

SUMANAS

The previous Web site illustrated the power and elegance that a Web site can have without any multimedia content. The Sumanas Web site (http://www.sumanasinc.com/index.html; Figure 4) is elegant and completely driven by multimedia animation, mostly implemented using Shockwave. Sumanas is a commercial provider of multimedia services, and they have made much of their product available for free on their Web site (smart marketing in my opinion). At www.sumanasinc.com/webcontent/anisamples/neurobiology/neurobiology.html they have assembled nine animated tutorials on neurobiology, most developed to supplement the 2004 textbook Biological Psychology by Rosenzwieg et al. (Sinauer).

These animations are well-drawn, moving illustrations emphasizing concepts, rather than photorealistic three-dimensional (3D) animations. Because the animations are fairly lengthy, running two to five minutes, the Shockwave downloads take one to two minutes, even on my fast Enterprise LAN connection. All the animations, covering Synaptic Transmission, Electrical Signaling in Neurons, AMPA and NMDA Receptors, Sound Transduction, Visual Pathways, Receptive Fields, Reflex Arcs, Receptors in the Skin, and PET, have a nice clean look. The accompanying narration sounds a bit stilted and crowded, almost like a wordy textbook figure legend. However, the narration can be turned off, and a step-through mode with explanatory text can be used instead, but I confess that I found it easier to follow the animations with the narration on.

The animation titled Sound Transduction (Figure 5) provides a good example of their animation style. The animation begins at the gross anatomical level and shows the viewer how sound waves progress through the outer, middle, and inner ear. At appropriate places, anatomical views are expanded to show physiological details. Once in the inner ear, the view is expanded to explain in detail the function of the cochlea’s basilar membrane. Viewers will see a simulation of the basilar membrane when stimulated by various sound frequencies. It is interesting to compare this very effective 2D style of animation to a similar 3D animation, called The Cochlea (Figure 5, right) on HHMI’s BioInteractive Web site (http://www.hhmi.org/biointeractive/neuroscience/animations.html). The 3D animation is perhaps more engaging in some ways, as is seeing the basilar membrane respond to music as opposed to pure tones. Nevertheless, the Sumanas animation is very easy to follow and goes on to explain the response of hair cells associated with the basilar membrane. Both styles of animation can play a role in illustrating difficult concepts.
Each Sunamas animation includes introductory and concluding text and a short interactive quiz. With only nine animations, the site cannot be comprehensive, but the topics are well chosen, covering molecular, cellular, and anatomical levels. While you're on the Sunamas Web site, you might also like to have a look at other animations that they've produced on topics such as stem cells and malaria (http://www.sumanasinc.com/scienceinfocus/scienceinfocus.html).

TUTIS VILIS

The level of Shockwave animation featured on the Sunamas Web site is probably beyond the skill level and time available for most faculty to achieve. However, Tutis Vilis in the Department of Physiology and Pharmacology at the University of Western Ontario has developed two Flash-based Web courses that show the sort of excellence that an individual faculty-developer can achieve in multimedia Web design (http://www.med.uwo.ca/neuroscience/vilis/courses.htm).

The Physiology of the Senses (http://www.med.uwo.ca/physiology/courses/sensesweb/) is a Web-based course for undergraduates that is used by students in more than 60 countries. The site is well thought-out and organized, with the end-user student in mind. It starts off by helping you check for proper Flash plug-in functionality. The primary learning modules are designed in Flash, but instructions and many materials are also provided in PDF format. Each of the 12 topics (including The Eye, Visual Cortex, Visual Perception, Association Cortex, Touch, Muscle Sense, Balance, and Memory) includes an animated Flash session, a PDF with excellent illustrations and explanations, a set of interactive “practice” problems, and a set of links related to the particular topic. All of the course information is organized in an easy-to-view grid, with each feature opening in its own window. Therefore, the entire course and user instructions can be organized in outline manner on a single homepage for the course.

Dr. Vilis’ other Web-based course is a neurophysiology course designed for first-year medical students (http://www.med.uwo.ca/physiology/courses/medsweb/). It is really a nine-part overview of the brain, emphasizing functional anatomy, but also including some cellular and molecular information. I really liked Lesson 5, Vestibular System and Eye Movements (Figure 6).

The lesson begins with a schematic of the ear and related organs. A sort of blue gut tube at the bottom of the page shows you where you are in the lesson and also doubles as navigation if you want to skip ahead or go back. The first several sections of the lesson cover gross anatomy and general functions of hearing and balance. Next, some details of how kinocilium function is presented and is nicely related to movements of the head. The semicircular canals get a similar treatment, and then the lesson moves on to the vestibulocular reflex. A very effective animation shows how each eye’s movements need to be linked to the other, as well as to movement of the head, in order to maintain a stable visual image on the retina. In a series of very nice steps, the entire circuit is laid out in cellular detail, always related to animated movements of the head and eyes. The feature ends with engaging visual explanations of nystagmus and dizziness.

In addition to excellent animation, the site makes good use of the vector-based graphics of Flash, meaning that the images and text scale to whatever size you make your browser window, a blessing to those of us with creeping presbyopia, a topic that I hope Dr. Vilis will add to his course soon.

OTHER NEUROSCIENCE WEB SITES IN BRIEF

The Mind Project (http://www.mind.ilstu.edu/curriculum/) focuses on the development of curriculum modules for teaching the cognitive sciences. The Project is looking for instructors to author and pilot modules.

Focus on Neuroanatomy

http://www.med.harvard.edu/AANLIB/home.html: The Harvard Brain Atlas is a repository of human brain images. Interactive tools allow you to advance serially through brain sections in various planes, viewing images produced by various imaging technologies.


http://www.brainmuseum.org/: The Brain Museum is an extensive online photo collection of mammalian brains, including information on development and function.

http://synapses.mcg.edu/: Synapse Web has excellent images and tutorials focusing on synaptic structure.

http://isc.temple.edu/neuroanatomy/lab/: The Neuroanatomy Lab is a good comprehensive neuroanatomy Web site with lots of images and interactive quizzes, emphasizing clinical aspects.

Focus on Neurology

http://www.advancesinneurology.com/: Advances in Neurology focuses on epilepsy, Parkinson’s disease, and Alzheimer’s disease. The many videos are aimed primarily at physicians and patients.
Focus on Kids

http://faculty.washington.edu/chudler/neurok.html: Eric Chudler’s Neuroscience for Kids has been a favorite of teachers since the mid-1990s. The site includes neuroscience in the news, hands-on student activities, and online interactive games.

http://www.brainsrule.com/: Brainsrule is designed to appeal to kids with interactive games and lots of cartoon graphics and includes teacher curriculum pages.

http://www.dana.org/brainweb/: The Dana Foundation’s Brainweb focuses on mental health, including brain injury, aging, sleep, addiction, and inherited diseases.

http://thalamus.wustl.edu/course/: Washington University School of Medicine Neuroscience Tutorial presents the basics of clinical neuroscience emphasizing neuroanatomy.

http://www.sci.uidaho.edu/med532/: The University of Idaho’s Nervous System Course is a very nicely laid out introduction to medical neuroscience.
The central nervous system (CNS) is the first adult organ system to appear during vertebrate development, and the process of its emergence is commonly called neurulation. Such biological “urgency” is perhaps not surprising given the structural and functional complexity of the CNS and the importance of neural function to adaptive behavior and individual survival.

Soon after an egg is fertilized, it is subdivided into many cells, and those cells are rearranged in a process called gastrulation, which begins the formation of the CNS. In fact, neurulation seems to begin before gastrulation is completed, and the two processes appear seamlessly joined, suggesting that many of the cells and underlying mechanisms are involved in both processes. Thus, students wishing to understand neurulation should begin their study with a review of gastrulation and its mechanisms using one of several text books (e.g., Gilbert, 2003) or an excellent monograph (Stern, 2004). They might also find useful a recent review of videos and published papers dealing with the subject (Watters, 2005).

NEURULATION IN XENOPUS

Amphibian embryos provide possibly the best reference material with which to begin a study of neurulation, if only because their early development has been extensively examined. In the Xenopus frog, the early external events are documented by an excellent time-lapse video (Keller, 2002; see also, Keller and Shook, 2004, http://www.gastrulation.org/Movie13_1.mov). The film depicts neurulation from a perspective above the dorsal surface of the embryo (Figure 1) and nicely illustrates both the continuity of neurulation and gastrulation (ending as the video begins) and the correlation of each process with anterior–posterior elongation of the embryo. A detailed tutorial on amphibian development, including additional videos from the Keller lab, is available at Jeff Hardin’s Web site (http://worms.zoology.wisc.edu/frogs/welcome.html).1 Hardin’s time-lapse movie and tutorial are eminently suitable for study by introductory biology students. Moreover, by telescoping a cellular process requiring several hours into a visual record lasting less than a minute, time-lapse videomicroscopy produces a dramatic, if artificial, pedagogical aid. Students, and their jaded instructors, cannot help but be impressed by the morphogenetic changes that characterize neurulation, and it’s been my experience that discussion of the process and its underlying mechanisms more readily arises after their viewing than from “chalk-talks” or consideration of more static diagrams. The events, however, are complex spatially as well as temporally, and the few published videos available document only a few of the intricacies of the process. Such videos likely raise more questions than can be readily answered.

In brief, during neurulation the CNS begins as a somewhat flattened layer of epithelial cells on the upper or dorsal surface of the amphibian embryo. This flattened layer is called the neural plate (Figure 1A) and appears with the completion of gastrulation. With the onset of neurulation, the embryo also begins to elongate along its anterior–posterior axis and become restricted in girth. As neurulation progresses, elongation continues, and the lateral edges of the plate, or neural folds, become prominent and begin to bend toward each other to form a tube (Figure 1B). As the neural folds of the plate close, the dorsal midline of the embryo begins to sink somewhat below the surface, and in making contact, the central depression in the plate becomes more distinctive (Figure 1C). As they become contiguous, the folds fuse (Figure 1D), producing a hollow neural tube and a continuous dorsal epithelium, which effectively displaces the neural tube into the interior of the embryo. After separation from the overlying epithelium, the neural tube thickens differentially as various regions begin to form different subregions of the brain and spinal cord.

1 Unfortunately, the vertebrate embryos most available for study are also relatively yolky and opaque. Correspondingly, internal features of gastrulation and neurulation are difficult to follow in living material. In this regard, another film of Xenopus development at the Hardin Web site (http://worms.zoology.wisc.edu/frogs/gastxen/gastxen_sagview.html) is noteworthy for its interior view of gastrulation along the sagittal plane (midline) of a living embryo.
To understand how neurulation comes about, students should be encouraged to review the Keller video at the Hardin Web site and to discuss the cellular biology of gastrulation and neurulation (cf. Keller and Davidson, 2004). Doing so will likely generate numerous questions. For example, to what extent might elongation of the embryonic axis result from the epibolic and involutional movements characteristic of amphibian gastrulation? Might these internal movements produce tension in the overlying surface epithelial layer that initiates lateral folding (cf. Keller, 2002)? How might changes in the shape, adhesion, and relative position of the epithelial cells, for example, contribute to the folding of the neural plate and, eventually, to the closure of the neural tube? Could the convergence of dorsal plate cells toward the midline and their intercalation also produce some of the axial extension observed in embryos at this stage of development, in a process called “convergent extension” (Keller and Davidson, 2004).2

In reviewing the Keller video and considering these questions, students may generate many questions of their own, more than can be currently answered. Fortunately, an earlier article by Davidson and Keller (1999) describes the medial movement of neural cells, their intercalation and convergent extension, and provides important clues and some answers to underlying mechanisms. Unfortunately, the videos accompanying this article will be difficult for most undergraduates to appreciate, because they depict the behavior of loosely associated cells, called neural crest cells, photographed at high magnification, under low light conditions, and without spatial context. Given these imaging limitations, it’s very difficult to visualize how their behavior is related to neural fold formation and closure, although the transverse sections examined by confocal microscopy provide excellent static images for temporal reconstruction (Davidson and Keller, 1999). More advanced students might enjoy discussing the results and the combined application of in situ RNA hybridization, confocal fluorescence microscopy, and videomicroscopy to correlate gene expression, cell movement, and neural tube formation.

Students who question how convergent extension and related cell movements might be coordinated will be interested in an article (with videos) describing the effect of Dishevelled (Dsh) on neural tube closure in Xenopus embryos (Wallingford and Harland, 2002). Figure 2 depicts the comparative development of embryos that had been injected with a mutant Dsh alongside embryos from the same population of fertilized eggs that had not been injected. All of the embryos began neurulation at about the same time, and all exhibited neural folds that became prominent and began to converge. The neural tubes of half of the injected embryos and all of the un.injected controls closed, although the rate of closure seemed faster in control embryos than in injected ones (compare the right-hand and middle panels in Figure 2). Half of the injected embryos, however, exhibited neural folds that failed to close (left-hand panel of embryos in Figure 2). Interestingly, the embryos with neural tubes that failed to close also seemed shorter and more stubby than the controls, and some students might immediately postulate a mutant disruption of convergent extension. To test this hypothesis, students could examine the sectioned material and other videos described in the article and discuss the details of mutant Dsh dysfunction. Other students might wonder how a mutant protein could exert partial effects, preventing closure in some embryos but not in others, and discussion of these results could lead to a fruitful exploration of 1) how “mutant effects” are produced using injection of mutant material into organisms such as Xenopus that lack collections of mutant strains, using injection of mutant material; 2) how mutant effects would occur in a “dominant negative” manner; and 3) how a spectrum of phenotypes might result from translational competition of the wild-type and mutant mes-

---

Footnote 2: Convergent extension is one of the more important developmental mechanisms underlying morphogenesis, and its three-dimensional features are difficult to illustrate to introductory students with static diagrams. Fortunately, an excellent article (Glickman et al., 2003) and 18 videos (http://dev.biologists.org/cgi/content/full/130/5/873/DC1) illustrate this process in wild-type and mutant zebrafish development. Unfortunately for the purposes of this Feature, however, the structure being examined and illustrated is the notochord, the embryonic forerunner of the vertebral column, which is not part of the CNS and which is derived from mesodermal cells that entered the embryo during gastrulation.
sage. (Although unstated by the authors, the injected material is likely mRNA [Sive et al., 2000].) Somewhat more critical students also might wonder whether saline-injected, rather than uninjected, embryos would have provided better, more reliable controls.

Students who have taken a course in developmental biology should recognize Dishevelled as a downstream signaling peptide in the Wingless-int (Wnt) pathway. They would thus be aware of this pathway’s important role in establishing tissue boundaries in Drosophila and Caenorhabditis elegans embryos, via activation of β-catenin, a putative transcription factor regulating cell division (cf. Gilbert, 2003). Because Dsh and β-catenin also help establish the dorsal axis of very young Xenopus embryos (Miller et al., 1999), and thereby affect the later development of such dorsal structures as the neural tube, some students might wonder how an injected, mutant form of Dsh could disrupt neurulation without apparently exerting a more global disruptive effect on axis formation. More thoughtful developmental biology students will recognize this paradox as another instance involving the timing of developmental events, where critical cues exert subtle effects well in advance of obvious phenotypic outcomes. Unfortunately, the authors do not indicate when mutant Dsh was injected into their Xenopus embryos, but useful discussion could be devoted to consideration of two hypothetical if/then scenarios. If injections occurred after the dorsal axis had been set, then the paradox is easily resolved and requires only a recognition that once axis formation is determined, subsequent developmental events along the axis are often inevitable and unchangeable. (Of course, such discussion quickly begs the more interesting question as to how exactly stable determination comes about and is maintained.) If, however, injection occurred right after fertilization, at the time the dorsal axis was being determined, then a different, but equally interesting, discussion would center on the developmental importance of maternal proteins and their localization in early Xenopus development as well as the timing of “new” mRNA synthesis. Cell biology students, however, might be more familiar with β-catenin’s role as a linker peptide anchoring cytoskeletal elements to integral plasma membrane proteins called integrins and as a regulator of cytoskeletal protein phosphorylation (Lodish et al., 2004). A fruitful dialogue between these two groups of students might well produce some additional hypotheses how Dsh could regulate neurulation, both through differential growth and tissue delimitation as well as through those polarized changes in cell shape and attachment thought to be important for convergent extension.

NEURULATION IN FISH

Neurulation in bony fish seems to occur in much the same way as it does in amphibians, but the neural folds are much less prominent and form a nontubular thickened tissue often called a neural keel (which appears wedge-shaped in cross-section). The keel gradually rounds up into a solid neural rod and not a neural tube. The neural rod then forms a hollow core and tubelike structure (Langland and Kimmel, 1997). Much less is known about this “secondary” form of neurulation, although the Wnt pathway and β-catenin are implicated in its regulation (Gilbert, 2003). An interesting discussion for students in an intermediate-level course in developmental biology could be devoted to exploring mechanistic variations between the two forms of neurulation.

Any study of neurulation in bony fishes would likely involve genetically tractable zebrafish embryos, which are easily raised in the laboratory and increasingly used for developmental studies. Students could begin their study with a paper (and set of videos) describing the contribution of polarized cell division to keel formation, which also provides a succinct, well-illustrated summary of neurulation (Geldmacher-Voss et al., 2003). These authors used zebrafish

Figure 2. A dorsal view of three neurulating Xenopus embryos between the midneural stage 16 (A) and the neural tube stage 21 (C). The embryo on the left is expressing Xdd1, an injected mutant form of Dishevelled and fails to close its tube (open-NT). The middle embryo is also expressing Xdd1 but successfully closes its neural tube (closed-NT) while lagging behind the control. The embryo on the right is an un.injected control. The arrows indicate neural folds in various degrees of closure. Although only relative developmental times are provided (in the form of “stages”), the same population of embryos was subdivided, with one receiving an injection of mutant Xdd1 and the other behaving as an un injected control; both sub-populations were reared under identical conditions. The time elapsed in A–C is approximately four hours. The entire video may be viewed at http://dev.biologists.org/content/vol129/issue24/images/data/5815/DC1/MOVIE1.mov.

NEURULATION IN FISH

Neurulation in bony fish seems to occur in much the same way as it does in amphibians, but the neural folds are much less prominent and form a nontubular thickened tissue often called a neural keel (which appears wedge-shaped in cross-section). The keel gradually rounds up into a solid neural rod and not a neural tube. The neural rod then forms a hollow core and tubelike structure (Langland and Kimmel, 1997). Much less is known about this “secondary” form of neurulation, although the Wnt pathway and β-catenin are implicated in its regulation (Gilbert, 2003). An interesting discussion for students in an intermediate-level course in developmental biology could be devoted to exploring mechanistic variations between the two forms of neurulation.

Any study of neurulation in bony fishes would likely involve genetically tractable zebrafish embryos, which are easily raised in the laboratory and increasingly used for developmental studies. Students could begin their study with a paper (and set of videos) describing the contribution of polarized cell division to keel formation, which also provides a succinct, well-illustrated summary of neurulation (Geldmacher-Voss et al., 2003). These authors used zebrafish
strains stably expressing histone2a chimeras fused with green fluorescent protein (GFP) to follow chromosome movements. To monitor the orientation of mitotic spindles during neurulation, they injected mRNA encoding a tau-GFP chimera; tau is a microtubule (MT)-associated protein (MAP) responsible for stabilizing bundles or fibers of MTs.

As the neural keel forms, cell divisions initially occur within the neural epithelium with mitotic spindles oriented at right angles to the neural plate. As daughter cells accumulate in a vertical direction, the keel thickens. The plane of cell divisions then rotates 90° such that in subsequent divisions, daughter cells are produced parallel with the surface and oriented at right angles to the developing anterior–posterior axis of the embryo. These features are illustrated in Figure 3, A–C, which was taken from a video of sequences lasting seven seconds, documenting ~102 minutes of development.

What is especially evident in the video (and Figure 3) is the extension of the neural plate along the anterior–posterior embryonic axis and concomitant shrinkage of the plate at right angles to this axis, which accompanies keel formation. This convergent extension of the keel is visually dramatic because greater amounts of tau:GFP chimeras are produced in cells along the left and, even more so, the right margins of the plate. The authors claim most cell divisions at this stage of neurulation are polarized (single arrows) and contribute to convergent extension of the keel. Students should be encouraged to assess this claim by mapping the orientations and mitotic fates of other apical cells and then discussing how such cells could extend the embryonic axis through their convergence at the midline. Some will likely agree that certain cell divisions seem to move daughter cells closer to the keel midline (single arrow, Figure 3, A and B), whereas others may argue that cytokinesis is producing daughter cells with other orientations in the apical layer and that these cells are being displaced toward the midline by different mechanisms. Curious students may wonder whether the brightly fluorescent cells on the right-hand side of the neural plate are somehow forming a “cleft” in the keel. To determine whether such a cleft is forming, they should be encouraged to examine other videos in the series and to speculate how such a cleft might be formed. More advanced students might also explore the article and other videos in greater detail, examining the localization of GFP chimeras with various cell junction markers. Further, they might consider what would be the effects of injecting antisense RNAs (“morpholinos”) for cell junction markers on both their localization and neural keel formation.

This article and others by Keller and his associates on Xenopus neurulation might also provide the basis for an extended journal club discussion of the cellular bases for neural morphogenesis.

As always, I welcome your comments about this Feature and your suggestions of other published, peer-reviewed papers and videos for future Features.

ACKNOWLEDGMENTS

Anna Strimaitis, a Middlebury undergraduate and Biology major, was immensely helpful searching the recent research literature for appropriate videos and preparing the still figures for this Feature.

REFERENCES


National and State Standards in Science and Their Potential Influence on Undergraduate Science Education

Jay B. Labov

Center for Education, National Research Council of the National Academies, Washington, DC 20001

INTRODUCTION

Standards in science and other subjects are a recent phenomenon in education, with most of them having been developed within the past 15 years. In 2002, Kimberly Tanner and Deborah Allen published in the first volume of Cell Biology Education (CBE) an important article about national science education standards for Grades K–12. That article likely provided for many readers of CBE their first glimpse of the standards movement in K–12 education and the potential impact of standards on higher education. The article focused on the two national science standards documents, Benchmarks for Science Literacy, published by the American Association for the Advancement of Science (AAAS; 1993), and the National Science Education Standards, published by the National Research Council (NRC; 1996), and what those standards specify about what students should know and be able to do at various grade levels in cell biology. Tanner and Allen (2002) also provided what have turned out to be some prescient ideas and questions about the more general roles of science standards in education in the United States.

Much has happened in K–12 education during the past four years and, for the following reasons, an update on the standards movement is warranted.

• Tanner and Allen reported how the national documents were designed to serve as guidelines for the development of state standards and pointed out some of the issues involved in translating the national standards to state frameworks. In the four years since Tanner and Allen published their article, all states except Iowa have developed their own standards. Indeed, some states are now in a cycle of revising their standards documents.

• During the past five years, the federal No Child Left Behind Act (NCLB) has become much more a part of the fabric of K–12 education. Most people who are familiar with this legislation identify it with new expectations for testing and accountability. What is less known is that states are also required to align standardized assessments and classroom instruction with their standards documents.

• To date, students have been tested only in reading and mathematics as part of NCLB. Science will be tested for the first time during the 2007–2008 school year. Thus, science will take on a much greater level of importance in public education in the ensuing years, and this increased attention could have important ramifications for higher education in the sciences.

This article addresses all of these issues and also provides readers who may not remember the details from Tanner and Allen with a more general and updated introduction to national and state science standards. It also provides a brief overview of the science content standards movement in the United States, insights into the forces that have caused the adoption of science standards to be contentious in some states, discussion of why college faculty need to learn more about their own state’s standards, and the roles that professional scientists might play in shaping those standards in the future. For a more comprehensive recent review of national and state K–12 science standards, see Sunal and Wright (2006).

A BRIEF HISTORY

Beginning with a historic conference of the National Governors Association and continuing into the early 1990s, President George H.W. Bush and the nation’s governors developed the National Education Goals (sometimes referred to...
Eight goals were articulated; goal 4 declared that by the year 2000, “U.S. Students will be first in the world in mathematics and science achievement.” To achieve this goal, the governors decided that national standards for science and other subjects should be developed. For science, the governors declared the following objectives.

Students in Grades K–12 will
Use scientific principles and processes appropriately in making personal decisions.
Experience the richness and excitement of knowing about and understanding the natural world.
Increase their economic productivity.
Engage intelligently in public discourse and debate about matters of scientific and technological concern.
Be aware of careers in science, technology, and the medical sciences.

These principles suggested approaches to science education that were very different from the prevailing teaching methods in several fundamental ways. First, instead of focusing almost exclusively on facts, these objectives also called for educating students to understand the connections between science and other types of knowledge and how science is relevant to their lives and their communities. Second, rather than emphasizing science education primarily for those students who were most likely to pursue careers in science or engineering (as had been promulgated in the post-Sputnik era), these objectives emphasized science education and scientific literacy for all students. Last, science was to be introduced to students much earlier in their academic preparation than was typical.

All of these changes had clear implications for the education and ongoing professional development of teachers, the numbers of teachers able to teach science, curriculum development and implementation, and even the physical spaces in which science would be taught. Clearly, some guidance was needed to help state education departments as well as local school districts and school personnel implement such sweeping changes in precollege science education.

This new perspective on science education was influenced greatly by the publication of Science for All Americans (Rutherford and Ahlgren, 1990) by AAAS. In response to both this publication and the directives of the National Governors Association, both AAAS and the NRC began work on producing national standards for science. AAAS’ Benchmarks for Science Literacy (Figure 1) were published in 1993 (AAAS, 1993) and focused on content standards.

The NRC released the National Science Education Standards (NSES) in 1996 (Figure 2). These standards deliberately embedded science content standards within a larger system of science education, thereby emphasizing that gains in student performance are dependent upon improvements in the entire system of science education and not solely on enhancements in content standards. Consequently, the NSES called for changes in six sectors of the education system that would be required to realize sustained improvements in student performance:

• Teaching
• Professional development for teachers
• Assessment
• Content
• Science education programs
• Science education systems

Both the Benchmarks and the NSES offer their content standards by grade bands rather than by individual grade levels; this gives schools, districts, and states flexibility in deciding when specific topics might be taught. Although the content standards in the Benchmarks and the NSES differ to some degree in emphasis (Table 1), an analysis by AAAS (1997) indicates that the Benchmarks and the NSES content standards are 90–95% congruent in their focus and subject matter. Thus, as individual states have adopted and adapted...
these national guidelines for their own, some have based their standards on one of these sets of national standards guidelines, whereas others have used aspects of both.

The NSES call for a very different way of presenting content and assessing students’ knowledge of science (Tables 2 and 3). The NSES view science education as something that students “do,” rather than something that is “done to them.” There is greater emphasis on integrating the processes and nature of science with content knowledge in the various scientific disciplines as a student progresses from the elementary through the secondary grades.

Not all content is of equal importance. Consequently, both the NSES and the Benchmarks stress that students will gain a deeper understanding and appreciation of science if they cover fewer topics and instead uncover some in greater depth, i.e., “less is more.”

The NSES also call for fundamental changes in what teachers should know and be able to do (Table 4), especially for elementary and middle school teachers who increasingly are becoming teachers of science. These recommendations suggest that new and very different approaches to teacher preparation and ongoing professional development are needed.

Since the publication of the Benchmarks and the NSES, both the AAAS and the NRC have published supplements to their original documents. AAAS has released several publications that focus on how to use the Benchmarks and the implications of their use in schools. An especially useful supplement is AAAS’ Atlas of Science Literacy (AAAS, 2001) that helps educators identify prerequisite knowledge and understanding in the content disciplines addressed by the Benchmarks that students need to study grade-level-appropriate material and to be prepared to progress to more advanced materials. The Benchmarks and all of these supplemental publications are available at http://www.project2061.org/.

Supplements to the NSES have focused on broader systems issues, including helping teachers understand the nature of inquiry (NRC, 2000), classroom assessment (NRC, 2001a), designing standards-based mathematics or science curricula (NRC, 1999), a framework for research efforts to investigate the efficacy of standards (NRC, 2001b), and a publication to help parents of school-age children and the general public understand the changes being promoted by the NSES (NRC, 1997).

**STATE-BASED IMPLEMENTATION OF NATIONAL STANDARDS**

**National Standards Are Not Federal Standards**

There are no mandated national standards for any subject in Grades K–12 in the United States. The responsibility for precollege education is vested constitutionally with state and local authorities. The Federal Government contributes approximately 8% of the total budget for K–12 education. Thus, the Benchmarks, NSES, and other national standards documents that were produced at the same time as, and subsequent to, these documents are intended to serve as guides that states can use to voluntarily develop and implement their own standards. However, when published the NSES represented a national consensus of the scientific and science education communities of what constitutes quality

<table>
<thead>
<tr>
<th>AAAS Benchmarks</th>
<th>NSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Nature of Science</td>
<td>Science and Technology</td>
</tr>
<tr>
<td>The Nature of Mathematics</td>
<td>Physical/Earth/Space Sciences</td>
</tr>
<tr>
<td>The Nature of Technology</td>
<td>Life Sciences</td>
</tr>
<tr>
<td>The Physical Setting</td>
<td>Science in Personal/Social Perspectives</td>
</tr>
<tr>
<td>The Living Environment</td>
<td></td>
</tr>
<tr>
<td>The Human Organism</td>
<td></td>
</tr>
<tr>
<td>Human Society</td>
<td></td>
</tr>
<tr>
<td>The Designed World</td>
<td></td>
</tr>
<tr>
<td>The Mathematical World</td>
<td></td>
</tr>
<tr>
<td>Historical Perspectives</td>
<td>History and Nature of Science</td>
</tr>
<tr>
<td>Common Themes</td>
<td>Unifying Concepts and Processes</td>
</tr>
<tr>
<td>Habits of Mind</td>
<td>Science as Inquiry</td>
</tr>
</tbody>
</table>

Topics are organized such that similar topics in each column are displayed on the same row.
Table 2. The NSES stress a changing emphasis on scientific content and process

<table>
<thead>
<tr>
<th>Less emphasis on</th>
<th>More emphasis on</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowing scientific facts and information</td>
<td>Understanding science processes and developing abilities of inquiry</td>
</tr>
<tr>
<td>Studying subject matter disciplines (e.g., physics, earth sciences) for their own sake</td>
<td>Learning subject matter disciplines in the context of inquiry, technology, science in personal and social perspectives, and history and nature of science</td>
</tr>
<tr>
<td>Separating science knowledge and science process</td>
<td>Integrating all aspects of science content</td>
</tr>
<tr>
<td>Covering many science topics</td>
<td>Studying a few fundamental science content</td>
</tr>
<tr>
<td>Implementing inquiry as a set of processes</td>
<td>Implementing inquiry as instructional strategies, abilities, and ideas to be learned</td>
</tr>
</tbody>
</table>

This information is from NRC (1996), p. 113.

Table 3. The NSES stress a changing emphasis on assessment of scientific knowledge and understanding

<table>
<thead>
<tr>
<th>Less emphasis on</th>
<th>More emphasis on</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessing what is easily measured</td>
<td>Assessing what is most highly valued</td>
</tr>
<tr>
<td>Assessing discrete knowledge</td>
<td>Assessing rich, well-structured knowledge</td>
</tr>
<tr>
<td>Assessing scientific knowledge</td>
<td>Assessing scientific understanding and reasoning</td>
</tr>
<tr>
<td>Assessing to learn what students do not know</td>
<td>Assessing to learn what students do understand</td>
</tr>
<tr>
<td>Assessing only achievement</td>
<td>Assessing achievement and opportunity to learn</td>
</tr>
<tr>
<td>End-of-term assessments by teachers</td>
<td>Students engaged in ongoing assessment of their work and that of others</td>
</tr>
<tr>
<td>Development of external assessments by measurement experts alone</td>
<td>Teachers involved in the development of external assessments</td>
</tr>
</tbody>
</table>

This information is from NRC (1996), p. 100.

science education and the educational systems needed to support that education. They were reviewed by thousands of scientists and science educators and by dozens of professional societies before their release.

State standards are now the predominant influence on K–12 education, and there is considerable variation from state to state in their use of the NSES and the Benchmarks and in their adjudged quality (e.g., Gross et al., 2005). As a specific example, evolution is a subject that has received considerable attention by the media, policy makers, and the public in both national standards documents. Some states have adopted these recommended standards faithfully, whereas others have eliminated selected components or do not mention evolution at all (Lerner, 2000; Gross et al., 2005). In other cases, there has been great controversy about the amount of content that students should be required to know and at what grade levels they are expected to know it. Political and other considerations continue to influence the state-based adoption process as individual states revise their standards every five to seven years.7

The proliferation of state standards has resulted in some unintended consequences. For example, science textbook publishers and curriculum developers who previously only had to show that their products were consistent with one or both national standards documents to be adopted now have to tailor their products to the many different state standards to be considered. Such pressures can lead to fragmentation of content or production of textbooks that respond to the “lowest common denominator.”

During the past few years, state standards also have taken on increasing prominence because of NCLB, which mandates that students be tested on content that is tied to a state’s standards in a particular discipline. Each state must administer tests aligned to those standards. However, since

Table 4. Excerpts from NSES of standards for the professional development of teachers of science

<table>
<thead>
<tr>
<th>Standard</th>
<th>Excerpt</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>The professional development of teachers of science requires learning science content through the perspectives and methods of inquiry . . .</td>
</tr>
<tr>
<td>B</td>
<td>Professional development of teachers of science requires integrating knowledge of science, learning, pedagogy and students, applying that understanding to science teaching . . .</td>
</tr>
<tr>
<td>C</td>
<td>The professional development of teachers of science enables them to build the knowledge, skills, and attitudes needed to engage in lifelong learning . . .</td>
</tr>
<tr>
<td>D</td>
<td>Pre-service and in-service professional development programs for teachers are coherent and integrated . . .</td>
</tr>
</tbody>
</table>

This information is excerpted and modified from NRC (1996).

---

7 See, for example, the controversy that has arisen in California at http://www.sci-ed-ga.org/standards/analysis.html.
the law’s inception, schools have only been held accountable for testing and demonstrating adequate yearly progress\(^8\) in reading and mathematics, and only for Grades 3–8 in these subjects. Science will be tested beginning in the 2007–2008 school year and then only in each grade band corresponding to the elementary, middle, and secondary grades.

There are three important consequences of this process that should concern scientists and science educators:

- In some districts teachers, especially in the elementary grades, have been asked to reduce or eliminate the teaching and learning of science to allow more time for preparation in reading and mathematics. Districts often do not embrace the interdisciplinary concept that science can serve as an effective vehicle for learning of mathematics and development of reading skills, both essential for successful performance in science. Consequently, reduction in class time doing science may result in a cohort of students who are ill-prepared to appreciate and succeed in science when testing does begin.
- When science is finally tested in 2007–2008, or when districts recognize that they have to begin preparing students for the science examinations, they will rely heavily, if not exclusively, on their state science standards, which often reinforce the learning of facts rather than the more systemic and integrative approaches that are emphasized by the NSES and the Benchmarks. Unless state standards both require and reinforce the notion that quality science education also includes exploration, data analysis, and developing deep conceptual understanding of topics, teachers will be under pressure to focus primarily on factual information, lower-level thinking skills, and limited conceptual understanding. Moreover, these conditions suggest that teachers and school administrators will tend to focus on the specific content and examples that are planned for the assessment.
- NCLB permits individual states to use any assessment instruments they wish as long as they align with that state’s content standards. However, assessments that authentically measure students’ deep conceptual understanding, their skill and ability to explore, transfer knowledge from one topic to another, and to synthesize and draw conclusions from data are more expensive and difficult to develop, administer, and score than tests that focus on factual knowledge. Thus, there likely will be strong financial pressure to use the less expensive, less rigorous instruments that are currently available. These kinds of assessments would send strong messages about the kind of science education that is valued and could reverse some of the gains that are beginning to be reported around the country. The NRC has published several reports that focus on these issues (NRC, 2001c, 2003, 2005).

Science education is at a crossroad and will continue to be over the next several years. Whether our nation will finally realize the kind of quality science education that was envisioned in the national standards documents, and is required for the United States to continue to lead the world in science and technology innovations and in the practical application of those innovations, remains to be seen. It is incumbent upon today’s scientists and science educators, especially those who are involved with higher education, to become more knowledgeable about the vision and emphases of the national standards and their reflection in the standards of the state in which they live and work, and the influential role scientists can play in shaping development and revisions in those standards and assessments.\(^9\)

College-level scientists can become engaged with these efforts in many ways. For example, they can work with state boards of education to review science standards in their specific disciplines or more broadly when those standards are being revised. They can work locally on selection committees for textbooks and other science education resources to help district leaders, administrators, and teachers determine whether those resources conform to state standards and meet the need for high-quality materials. As parents, grandparents, and citizens, scientists can play important roles in their schools and communities as advocates for high-quality science teaching and learning for all students.

The focus, quality, and effectiveness of today’s science education programs ultimately will result in college students who are or are not prepared to engage in science at the postsecondary level. Thus, postsecondary educators have a vital long-term stake and self-interest in today’s K–12 science education programs and the policies that govern them. Higher education has had little influence to date on the development or implementation of standards because scientists have not been well represented at the table at either the district or state levels. For all the reasons outlined above, the time to do so is now.

---

\(^8\) For more information from the U.S. Department of Education about this component of the law, see http://www.ed.gov/nclb/accountability/ayp/yearly.html.

\(^9\) Access to individual state standards is available through a state’s Department of Education or through http://www.education-world.com/standards/state/index.shtml.

**REFERENCES**


Of the dozens of books about teaching and learning that I have read over the past decade, none has reverberated in my thoughts like Learner-Centered Teaching. This book confronted me with the great distance between my personal definition and practice of "learner-centered teaching" and the potential of a deeper, more authentic learner-centered approach. It confronted me with the lapses between my "talk" and my "walk" when it comes to attitudes toward students, course design, expectations, power, and many other aspects of teaching. It helped me think more creatively about helping students with diverse learning styles and needs to thrive in my classes. What makes this book so valuable and thought provoking?

Weimer begins her book with a discussion of her redesign of a communications course to be more learner centered. In this new course, students were provided with a menu of assessment options, so that specific decisions about how they would be evaluated were determined by the student herself or himself. Based on my own teaching experience, the initial response of Weimer’s students was predictable: They were stunned and uncomfortable, looking for the "catch." They could not initially believe that they did not have to take the exams, but could earn their points in other ways. The ultimate response of her students was less predictable but not surprising in hindsight: Students became more engaged with the material and worked harder than they did in Weimer’s previous teacher-centered course. This anecdote, told in a pleasant conversational style, leaves the reader with a strong desire to design and teach similar courses, courses in which students find the content interesting and relevant, are animated in class discussions, and learn more effectively.

The course itself is also very interesting as a model for creative learning activities. For example, students in the class could earn points by “class participation.” Each class participant was paired with another participant, with each member acting as a participation “coach” for the other. The class as a whole defined what class participation meant, and the participants earned their points by writing three papers: a paper outlining their participation goals in the context of the class definitions, a feedback paper to their partner about his or her class participation, and a self-assessment of his or her own participation. Such an assignment assessment plan involving student coaches was delightfully new to me and made me consider other ways in which my students might participate in feedback to one another.

In the next five chapters, Weimer presents five changes that she believes are necessary to create an authentic learner-centered teaching practice: 1) shifting the balance of power more toward the student; 2) using (rather than covering) content as the means to achieve higher-order learning goals; 3) changing the teacher’s role from that of telling/doing to that of designing/modeling; 4) helping students accept the responsibility for learning; and 5) adapting the purpose and processes of evaluation to promote learning. Maintaining the highly readable, conversational writing style that draws you in at the beginning, Weimer elaborates on each of these changes in a separate chapter. Each chapter ends with a very effective summary paragraph that reinforces the key messages of the chapter. Because the chapters are short and focused and the key messages are restated at the end, the reader falls almost automatically into a pattern of intense reading, followed by reflection. In my case, the book was read over a couple of weeks of commutes in the "Campus Connector" between Minneapolis and St. Paul. It wasn’t that the book could not be read at one sitting. Instead, after reading each chapter, I found myself needing to internalize the information and to envision the changes that I might make to my own teaching. How should I structure my new seminar to give students more decision-making power? How could I help students assume greater responsibility for their learning? How could I design assessments that help students actually master the course content? How would students respond to these changes?

Although the first part of the book is essentially a detailed explanation of learner-centered teaching, the second part deals with how to implement the necessary changes to achieve that goal. Surprisingly, the implementation part of
the book was less valuable than the definition part. For example, as I’m sure the readers of CBE—Life Sciences Education recognize, changes such as those recommended above are likely to cause resistance from every quarter, including from both students and faculty colleagues. Weimer discusses where the resistance arises. For example, students resist learner-centered approaches because these approaches require more work, are more threatening than traditional teacher-centered approaches, force students to take responsibility for their learning, and may be difficult for some students. Weimer’s answer to this resistance is communication with students. Thus, although the analysis of the source of resistance was valuable, the solutions appeared superficial. Nevertheless, this portion of the book forces faculty who are considering moving toward learning-centered teaching to understand some of the challenges that face them.

After reading this book and pondering on its message for several months, I returned to it recently and revisited key points. The ideas and conversational style still resonate with great power: the function of content as a vehicle for skill building; the teacher as midwife, trail guide, or coach; the roles of self-assessment, formative assessment, and grades. Powerful ideas. Difficult ideas. Revolutionary ideas.

I reconsidered my own attitudes toward teaching: I envision my courses to epitomize the “guide on the side” style of teaching. I pat myself on the back as I include bits of active learning exercises or “clicker questions” at the midpoint of the 30 PowerPoint slides that I need to cover in that day’s lecture. I’m proud of the tight organization of our courses that provide students with clear expectations from the time they enter the course.

Although these changes are authentic improvements since my first dismal teaching experiences (I’m still profoundly sorry for those first students!), Learner-Centered Teaching puts these efforts into an intense new light. It reveals that I still have a long way to go! Few books on education have had the impact of Learner-Centered Teaching on my thinking about how to be a better teacher. And, although I’m not yet truly walking the walk, I think I’m at least up on all fours and moving ahead.
Review

Moving from Medieval Apprenticeships to Reflective Practice


Reviewed by Keith Garbutt, Department of Biology, West Virginia University

INTRODUCTION

The university is an institution with its roots in the medieval period. For example, the gowns and regalia we don for ceremonies such as convocation and graduation hearken back to the robes worn by the clerics who founded the early universities. In spite of our proud preservation of such centuries-old traditions, we still like to think that our modern institutions have progressed somewhat from the early days of Padua and Oxford. However, in at least one critical aspect, we continue to operate in recognizably medieval conventions. The training of the next generation of the professoriate remains a process that reflects the apprentice system of the guilds of the Middle Ages. In any reasonably sized laboratory, we find undergraduate and graduate students who could quite easily be considered junior and senior apprentices. Postdocs correspond to medieval journeymen, and, of course, the professor in charge of the group would have been known as the “master” in the guild system. Although this apprentice model has admirably served those of us who have successfully made it to the professoriate/master level, it also has left us with some serious deficiencies in training faculty, especially in the areas of teaching and mentoring.

At most institutions, the apprenticeship programs we currently use to train our students in the elements of teaching and mentoring are a process of modeling. If the model is good—if the master provides a living example of excellent teaching and mentoring to the apprentices—then we have a reasonable chance of turning out individuals who themselves will be good teachers and good mentors. But, if the model proves to be an indifferent or poor teacher or mentor, then we will be turning out a new generation of scientists who are equally indifferent or poor teachers and mentors. This modeling system is analogous to the ways in which most of us approach parenting; we simply apply the skills, attitudes, and approaches we absorbed from our caregiver models to our children. However, as we grow older, we all discover, to our horror, that we are turning into our parents.

We did not only learn and apply the effective ways in which we were guided and nurtured but also the flawed approaches and perspectives. I fear that many of us are also discovering, to our equal horror, that we are turning into our advisors.

In many research programs, the individuals who bear responsibility for the day-to-day activities of the undergraduates in a laboratory are the graduate students and the postdocs. Thus, these senior apprentices and journeymen would benefit from learning the skills and approaches needed to be effective mentors for these undergraduates. Conversely, the undergraduates would have a more rich and successful experience in our laboratories under these skilled mentors. Entering Mentoring: A Seminar to Train a New Generation of Scientists, from the Wisconsin Program for Scientific Teaching, was developed for this purpose and aims to break the cycle of developing mentoring skills exclusively based on the models provided by one’s own mentors. The authors make some fundamental assumptions about the situation in which mentoring will take place: it was designed for, and works exceptionally well for, a life sciences program that has a summer research experience for undergraduates. However, our experience has shown that it also works in a regular semester setting and for graduate students from a wide range of disciplines.

Entering Mentoring provides a detailed plan for conducting a seminar that helps graduate students and postdocs reflect seriously upon the process of mentoring research students. Because the program is set up primarily with a summer undergraduate research experience in mind, it includes plans for a series of eight weekly sessions that, concurrently with their work with the students in the summer program, lead the graduate student or postdoc mentors through the intellectual issues, technical issues, personal growth issues, and interpersonal issues associated with effective mentoring. The concurrency of the seminar with mentoring practice provides richness and relevance to the theory and information presented in the seminar. As a result, the undergraduate “mentees” immediately benefit from the improving mentoring skills of their graduate student or postdoc mentor.
In considering the feasibility of offering a seminar during the summer, one issue that certainly raises its head is whether we can afford the time to do such a thing. Having now used this manual on two occasions, my response would be, “Not only can we afford the time, but also we absolutely must afford the time if we are going to make the experience for undergraduates in our laboratories a truly exceptional one.” In addition, because of the rich background provided in the manual, the time needed for the seminar leader to prepare is relatively modest. Although I would not agree with a professor who claimed that it was possible to prepare for this course in the elevator on the way to the seminar, certainly it is not necessary to put in many hours of preparation before leading the seminar. The materials presented in the book and some supplementary readings prepare one well to lead a quality seminar.

The description of each session is split into facilitator’s notes and materials for the mentors, including readings on topics such as scientific teaching, the role of mentor, the challenges and benefits of diversity, and, worthy of special mention, an excellent essay by Jo Handelsman: “Righting, Writing.” Each weekly session leads the participants through the process of mentoring, initially beginning with establishing a good relationship and developing a philosophy for mentoring. It then covers topics of communication, goals, and expectations. One session is devoted to evaluating the progress of both the mentors and their mentees, and the manual provides evaluation protocols for the mentee, the mentor, and the facilitator of the program. Other sessions are devoted to applying the wisdom of the group to help mentors who face situations that are proving to be challenging or troublesome. The final session revolves around preparation and discussion of the formal “mentoring philosophy” of each of the mentors, including how this statement might be viewed by a search committee.

We have now offered the Entering Mentoring seminar at my institution on two occasions: once during the regular semester and once over the summer. On both occasions, the seminar was offered for credit. At the first offering, students in the Department of Biology, my home department, were encouraged to take the seminar if they were working in a laboratory with undergraduate students, but it was not mandated. Additional individuals attended from other departments, in response to an e-mail sent out by the Office of Graduate Studies. This inaugural semester session included 15 students. The summer session was linked to an EPSCOR Research Experience for Undergraduates (REU) Program award.

In fact, the reviewers of the proposal commented positively on the inclusion of the mentor training in the proposal evaluation. Faculty who wished to take part in the REU program were required to assign a graduate student or postdoc as a mentor, and the mentor was required to attend the Entering Mentoring seminar. Perhaps surprisingly, faculty were very supportive, and a mentor was assigned from all 23 laboratories, although a couple of the mentors were not initially as engaged as the participants had been in the first offering when all the mentors were volunteers.

In both the semester and the summer incarnations, the program worked well. However, as with any “course in a box,” there are ways in which it can be modified and improved to fit the needs of any particular institution. During the first iteration, we discovered that some issues arose regarding concepts of ethics and fairness in the research environment. Thus, we modified the program in our second iteration to include an extra pair of sessions on basic ideas of ethical behavior within a research environment. Based on these sessions, we also gave a similar ethics seminar to the incoming undergraduate research students, many of whom seemed not to have had any training in this area.

I am sure that as the Entering Mentoring seminar is used at other institutions, others will find elements that they would like to add to the program. The beauty of its structure as only eight sessions means that the addition of extra modules is usually quite straightforward, particularly if one is offering the seminar during the regular semester rather than during a summer session. Additionally, elements of the manual can be used outside the seminar. For example, I used the “Righting, Writing” essay this semester with an incoming class of freshmen. It was exactly the tool I needed to help these students begin to understand the procedures of writing appropriately in the biological sciences.

Overall, Entering Mentoring is an excellent manual for taking an important step in moving from a medieval apprentice model to a more purposeful, reflective model for learning how to be an effective mentor. It provides a solid framework and foundation to explicitly teach the concepts of mentoring to our graduate students and postdocs, hopefully helping them be better mentors than we are ourselves.
Feature

Points of View: On the Implications of Neuroscience Research for Science Teaching and Learning: Are There Any?

A Skeptical Theme and Variations: The Primacy of Psychology in the Science of Learning

John T. Bruer
James S. McDonnell Foundation, St. Louis, MO 63117

NOTE FROM THE EDITOR

Points of View is a series designed to address issues faced by many people within the life sciences educational realm. We present differing points of view back to back on a given topic to stimulate thought and dialogue.

The focus of all contributed features and research articles in Issues in Neuroscience Education is the teaching and learning of neuroscience, from elementary school to graduate school audiences. However, neuroscience is unique as a branch of biology in that it includes the study of neuronal and brain mechanisms that may underlie learning. To highlight this unique position of neuroscience, we have chosen to focus this issue’s Points of View on how research findings in the field of neuroscience may or may not have implications for the teaching and learning of science in general. We invited authors to address the following questions:

• What are the current implications of neuroscience research, if any, for how to improve K–25 science teaching and learning in schools and universities?
• To what extent will neuroscience research into biological mechanisms of learning, memory, attention, and other brain functions inform educational practices and science teaching in the future?

INTRODUCTION

I am notorious for my skepticism about what neuroscience can currently offer to education. My skepticism derives from several concerns, but a common theme runs through all of them: attempts to link neuroscience with education pay insufficient attention to psychology. In what follows, I will present four variations on this theme. First, for those who are committed to developing a science-based pedagogy and solving existing instructional problems, cognitive psychology offers a mother-lore of still largely untapped knowledge. Second, attempts to link developmental neurobiology to brain development and education ignore, or are inconsistent with, what cognitive psychology tells us about teaching and learning. Third, cognitive neuroscience is the brain-based discipline that is most likely to generate educationally relevant insights, but cognitive neuroscience presupposes cognitive psychology and, to date, rarely constrains existing cognitive models. And fourth, the methods of cellular and molecular neuroscience are powerful, but it is not always clear that the concepts of learning and memory used by neuroscientists are the same as those used by psychologists, let alone by classroom teachers.

Cognitive Psychology: A Basic Science of Learning

My notoriety as a neuroscience and education skeptic derives primarily from my 1997 article “Education and the Brain: A Bridge Too Far” (Bruer, 1997). I argued that cognitive psychology, not neuroscience, is the strongest current candidate for a basic science of learning. The psychological research that appears most relevant to improved educational practice is research on human problem solving and expertise, a research program initiated by Newell and Simon (1972), and research on memory and knowledge organization as reviewed by Brown et al. (1983). Research in this tradition attempts to explain problem solving and learning by formulating and testing cognitive models that assume the existence of mental representations and functions. In the educational context, such models can describe novice behavior in a problem-solving domain, as well as expert behavior. Such models can guide learning by suggesting instructional interventions that might transform novice models into more expert models (Bruer, 1993).

In 1986, at the James S. McDonnell Foundation, I initiated a program, Cognitive Studies for Educational Practice (CSEP), that funded applications of cognitive principles to K–12 instruction (McGilly, 1994). That program provided
research support for collaborations between cognitive psychologists and classroom teachers to develop research-based methods for solving recognized instructional problems, that is, classroom problems teachers identified as significant barriers to learning. Among the issues teachers identified were students’ inability to read with understanding, to master elementary arithmetic, to understand fractions, to make the transition from arithmetic to algebra, to transfer scientific understanding from classroom to real-world problems, to learn American history, and to write a coherent essay. The psychologist–teacher teams developed instructional methods based on cognitive research to address these problems and tested their interventions in classrooms, often with educationally significant results. Research funded under the CSEP program, along with other cognitively inspired educational research, provide the substance for a series of reports published by the National Research Council (NRC). Neuroscience does not figure prominently in these reports even though brain scientists served on the boards and advisory groups that prepared these documents.

For example, How Students Learn (NRC, 2005) makes no mention of current or future implications of neuroscience for educational practice. Another NRC study, Knowing What Students Know (NRC, 2001), discusses the implications of brain science in a six-page appendix within the 315-page report. The report dismisses the folk-theoretic views about brain lateralization and learning, the advisability of extrapolating from effects of environmental enrichment on rodent brain educational practice, and the educational significance of critical periods. While concluding that applications of brain science to general education are currently limited, the report points to special education as a promising area for future applications, citing Michael Merzenich and Paula Tallal’s work on dyslexia (p. 109). The NRC study How People Learn: Brain, Mind, Experience, and School (NRC, 1999), a 300-page book, has one 10-page chapter on the brain. The study modestly concludes that brain research has established that structural changes in the brain encode learning and acknowledges that in the future, neuroscience might provide some practical benefits to educators.

So, I am not alone in my skepticism, or at least my reservations, about the relevance of current neuroscience to educational practice. There is a substantial research community within psychology—whose work focuses on teaching, learning, real-world instructional problems, and solutions to these problems—that shares my concerns. Given the problems that confront teachers in the classroom (as illustrated by the list above), as a practical matter, many of us see neuroscience as having currently little to contribute toward solving those problems. What is frustrating to educational researchers and others committed to developing a science of learning is that educators’ current fascination with synapses and brain images causes them to overlook a substantial body of psychological and behavioral research that could have immediate impact in the classroom.

**Synapses, Critical Periods, and Developmental Neurobiology**

What accounts for the current fascination with synapses, brain images, and learning? Although educators have always been intrigued by brain science (witness the long history of right- vs. left-brain learning in education and the media), it was the concerted effort of policy advocates in the mid-1990s, arguing for an expanded Head Start Program, that brought brain science and education to the covers of *Time* and *Newsweek* (Bruer, 1999). A policy argument intended to advance the legislative cause of the educationally disadvantaged, once it hit the newsstands, resonated with middle-class parents throughout the industrialized world and provided grounds for purveyors of educational materials and advice to develop and promote “brain-based curricula” and “brain-compatible learning” programs.

The brain and early childhood education argument was based on three well-established results from developmental neurobiology. First, in early childhood there are periods of rapid developmental synaptogenesis, followed by synaptic pruning. Second, there are critical periods in development when normal experience is required for normal development. Third, rodent studies have shown that rearing the animals in complex environments has demonstrable effects on brain structure. These results are spun into the following story: periods of peak synaptic density or high resting brain metabolic rate are the periods when children learn everything most easily, when experiences hardwire the brain for life, and when learning results in life-long neural changes. According to this literature, the period of peak synaptic density is the critical period for brain development. The more synapses that are used during this period, the more will be retained into adulthood, and more synapses equal higher intelligence. Parenting and education are exercises in synaptic conservation, where more and earlier experience saves neural connections, builds better brains, and maximizes intelligence. Depending on the author, the policy recommendation, or the educational program being peddled this critical period can be birth to three, birth to 10, or three to 12 years of age.

Of course, on closer consideration, the findings of developmental neurobiology do not cohere with this popular story. Synapse elimination is essential for normal brain development, and elimination is primarily under genetic, not environmental control. Critical periods do not always neatly correlate with peak synaptic density or resting brain metabolism, and critical period effects appear to be confined to acquiring species-general traits (e.g., vision, first language) and do not generalize to culturally specific learning, the kinds of learning that occur in school and in an individual’s daily life. Rearing rodents in complex environments affects brain structure throughout the life span, not just early in life when critical periods typically occur. More significantly, the popular “brain and education story” does not cohere with what experimental psychologists know about learning over the life span. Specifically, the proponents of the story ignore decades of psychological research that shows that rate and ease of subject matter learning depend more on prior background knowledge than on biological maturation and chronological age. The NRC reports abundantly document this finding.

Of course, this argument is aided and abetted by possibly overzealous neuroscientists who want to see their work as having implications for solving societal problems or who, concluding their published articles, speculate about the direction of further research. I have provided examples of this
difficulty elsewhere (Bruer, 2002). In the present context, one example will suffice.

In spring 2000, the popular press and brain-based educators became enthralled by the claim that the teenage brain was not yet mature. The hard science behind this claim was a paper reporting the results of a magnetic resonance imaging (MRI) study showing an increase in cortical gray matter in regions of the adolescent brain. The authors of this study speculated that the observed changes might be related to a second wave of synaptic overproduction: “It may herald a critical stage in development when the environment or activities of the teenager may guide selective synapse elimination during adolescence” (Giedd et al., 1999). The original study’s speculation about a possible second wave of synaptogenesis was highlighted in a U.S. National Institute of Mental Health press release (see NIH Publication No. 01–4929NIMH Press release, Child and adolescent mental health information at http://www.nimh.nih.gov/publicat/childmenu.cfm). The press release, while acknowledging that the cause of the changes in gray matter were not yet known, said that changes in the teenage brain may be due to the same “use-it-or-lose-it” principle that governs early development of the visual system. Synapses that get exercised are retained, whereas those that are not wither. Among brain enthusiasts, possible late synaptogenesis would extend the biologically privileged learning period into the adolescent years. The press release speculated about the implications of the researchers’ speculation.

This message became the scientific centerpiece for the May 2, 2000, White House Conference on Teenagers. On National Public Radio’s (NPR) Morning Edition coverage of the conference, “use it or lose it” was presented as science’s best guess of what was happening in the teenage brain: “If children are using their brain at this point for academics or sports or music or video games that is what their brain will be hardwired or optimized for” (www.npr.org/opt/collections/torched/me/data_me/seg_73624.htm provides the audio of the segment). This is now a third-order speculation.

The “spun” version also resonated with educators. The morning the NPR story ran I received the following e-mail from a teacher: “I heard this incredible piece on NPR this morning the abstract for which I will reproduce below. This has unbelievable developmental implications—helps explain why junior high school kids don’t learn anything! If the pruning of the brain actually happens twice, this also helps explain the incredible leap in learning rates of adolescents (since the pruning begins, not during the explosion of cell growth).”

Third-order speculation should not be the process by which neuroscience is perceived to have implications for education.

How can neuroscientists help educators now? It is possible that neuroscientists are not aware of how their work and their forward-looking, speculative hypotheses are perceived and interpreted within the educational and lay communities. They should think critically about how their research is presented to educators and the public and should avoid even the most innocent speculation about the practical significance of basic research. They should remind the interested public that we are just at the beginning of our scientific inquiry into how neural structures implement mental functions and how mental functions guide behavior. We neuroscience and education skeptics make this recommendation: “Neuroscience has advanced to the point where it is time to think critically about the form in which research information is made available to educators so that it is interpreted appropriately for practice—identifying which research findings are ready for implementation and which are not” (NRC, 1999, p. 114). We would also add that speculations about practical applications are not research information.

Cognitive Models Matter in Cognitive Neuroscience Too

In my 1997 article, I argued that, although cognitive psychology provides the best current and midterm future foundation for a science of learning, if neuroscience were to become relevant to education, it was most likely to be via its subdiscipline of cognitive neuroscience. But even here, the road to better learning is probably not as independent of psychological and behavioral science as some cognitive neuroscientists and educators might believe.

One of the impediments to cross-disciplinary dialogues about the relevance of science, or a science, to learning is a clear understanding of what the basic assumptions and methods of the various disciplines are and the levels of analysis at which they operate. Traditionally, psychology is best understood as a discipline that studies individual behavior. For most of its history, it has been a behavioral science in at least two senses. First, psychologists attempt to develop theories that explain behavior. Second, the data psychologists collect to frame and test their theories have been behavioral data—reaction times, eye movements, and number of items successfully remembered from a list. In North America, from the turn of the century until the mid-1950s, behaviorism dominated psychology. Behaviorists believed that only observable entities should be allowed into psychological theories and that all human behavior could be accounted for by chains of stimuli and responses, with no need to posit unobservable mental functions and concepts. In the mid-1950s, North American psychology underwent a “cognitive revolution,” wherein psychologists recognized that any adequate theory of human behavior required positing mental constructs and functions that were not directly observable (see Bruer, 1993). Psychology became a science of mind. Psychologists viewed the human mind as a computing device that contained both programs and data structures. Cognitive psychologists still used behavioral data, but now used it to frame and test hypotheses about what programs and data structures enabled human behavior. Cognitive neuroscience emerged as a new discipline in the early 1980s. Cognitive neuroscientists work at the interface of biological and behavioral science. Using both behavioral and biological measures of brain activity (single-cell recording, evoked response potentials, and brain-imaging technologies), cognitive neuroscientists attempt to discover the neural hardware that runs the mental software posited by cognitive psychological research. As we will see, one of the key issues in developing a coherent science of learning is being clear about these various levels of analysis and how these levels interact.

Compared with speculations about synapse formation, contemplating the educational implications of cognitive neuroscience is a welcome step in the right direction. The
reason for this is that cognitive neuroscience and an applied science of learning meet at an appropriate level of analysis. Cognitive neuroscience presupposes cognitive models. Furthermore, our current applied science of learning has established how cognitive models can contribute to improved teaching and learning, as we see in the NRC reports. Unlike cognitive psychology, which leaves educators and the public cold, cognitive neuroscience has boundless, albeit superficial, popular appeal. This popularity springs from its chief research tool: brain imaging. Cognitive neuroscientists present their data in colorful images, where highly active brain areas appear as bright patches within the brain.

Cognitive neuroscience is a hybrid discipline, a melding of cognitive psychology, systems neuroscience, and computational modeling. From the outset, the goal of cognitive neuroscience has been to identify neural structures that implement cognitive functions. In its initial decade, evidence for localization claims came primarily from lesion or electroencephalograph studies. One might say that cognitive neuroscience ceased to be a “boutique discipline” with the development of positron emission tomography and later functional MRI technology.

These advances in brain imaging technology allowed cognitive neuroscientists to gather data on brain activations in normal subjects. In 1988, Posner et al. (1988) articulated a research strategy for cognitive neuroscience, brain-imaging studies. The first sentence of their abstract stated: “The human brain localizes mental operations of the kind posited by cognitive theories.” Performance studies typical of psychological research, they argued, provide at best indirect and inconclusive evidence about localization of cognitive processes. Imaging studies provided a new source of direct evidence that allowed cognitive neuroscientists to test hypotheses about the localization of cognitive processes. That first sentence formulated the working hypothesis of cognitive neuroscientific brain-imaging research. The utility and power of that hypothesis is evident in the progress cognitive neuroscience has made to date.

That the working hypothesis remains central to cognitive neuroscience is evident in recent discussions of criteria that imaging studies should fulfill to be considered publishable. The editors of Nature Neuroscience (Editorial, 2001) suggested that among these criteria are the requirement that the study be hypothesis driven; that the study allow scientists to ask questions about basic cognitive processes, rather than identifying networks of brain regions activated by a series of tasks; and that the study include rigorous behavioral designs that ensure that the authors have isolated the cognitive process of interest.

Imaging studies allow cognitive neuroscientists to localize the processes, functions, and representations that cognitive psychologists have identified using methods of experimental psychology, cognitive constructs that have prior experimental support in behavioral and performance data. Cognitive neuroscience thus presupposes cognitive psychology. Cognitive neuroscience makes advances by providing better localizations of cognitive functions, but one could certainly argue that fundamental progress in cognitive neuroscience depends not only on the ability to identify cognitive processes and localize them, but also on the ability to analyze further these processes into their subcomponents at even higher levels of resolution. If the challenge is to understand at deeper levels the actual mental operations implemented in brain areas, then the cognitive models used in imaging studies must be continually refined (Posner and Raichle, 1994). Cognitive models are as fundamental to cognitive neuroscience as they are to our applied science of learning.

Reading instruction and treating dyslexia are areas of considerable educational import. Imaging studies on reading and dyslexia are cited as examples where neuroscience, in the form of cognitive neuroscience, is also thought to have implications for education. Among cognitive neuroscientists and educators, the imaging studies published by the Shaywitz Laboratory at Yale University are probably the best known. In a series of studies, these researchers have shown that there is a functional disruption of brain organization in adult dyslexics (Shaywitz et al., 1998), that similar disruptions are evident in posterior brain systems of child dyslexics (Shaywitz et al., 2002), and that a phonologically or phonics-based reading intervention results in the development of left occipitotemporal brain systems required for skilled reading (Shaywitz et al., 2004).

These imaging studies, as the authors acknowledge in their cited references, presuppose a long history of psychological research on reading and an even longer one in clinical neurolgy on dyslexia. The earliest theories of word reading and dyslexia derived from clinical neurological case studies in the late 19th century. These theories posited the existence of a sensory language center and a motor language center in the brain connected by a transmission pathway (Geschwind, 1979). Damage to either of the areas or the pathway resulted in different forms of dyslexia. Until the early 1980s, neuropsychologists’ primary concern was correlating disorders in cognitive functions, like dyslexia, with specific brain lesions. In 1982 Coltheart (1982) argued for a different approach. Neuropsychologists should engage in model building. Their goal should be to interpret patterns of impaired and preserved cognitive functions (dyslexia) in terms of an explicit model of the normal operation of those functions (word recognition). This approach led to the emergence of cognitive neuropsychology, wherein cognitive models of normal cognitive function provided a theoretical foundation for explaining cognitive deficits (Shallice, 1988).

Research on reading and dyslexia progressed on the basis of cognitive models of word recognition.

These cognitive models posit modules or computational processes that identify some visual stimuli as legal strings of letters according to the spelling rules of a language (orthographic representations), convert these strings of visually presented letters into sound patterns of the language (phonological representations), associate meanings with these sound patterns (semantic representations), and generate the motor programs needed to pronounce the words. Specific cognitive models differ in how these modules might be interconnected and about how the modules are implemented computationally (Coltheart et al., 1993; Harm and Seidenberg, 2004).

This cognitive research has supported a growing consensus that phonological processing is fundamental to skilled reading and that phonological processing deficits account for the most prevalent form of reading disability, phonological dyslexia. A fundamental assumption that has guided this research over the past 25 years is that theories of dyslexia should be grounded in our understanding of the psy-
chological mechanisms, i.e., the cognitive models, that support word recognition.

One motivation of the 1998 Shaywitz study was to determine why previous imaging studies that had attempted to identify a neural signature of phonological dyslexia had been inconclusive. Their solution was to develop a task hierarchy for an imaging study that would systematically tap components of the prevailing cognitive model of word recognition. These models predict that the most reliable indicator of phonological dyslexia is a reduced capacity to read pseudo- or nonwords (Castles and Coltheart, 1993; Stanovich et al., 1997). According to the theory, pronouncing such orthographically legal, but fictive, words (e.g., “mard” in English) places the highest demand on phonological processing. In the imaging studies, brain activations of dyslexics versus typical readers differ most from the control condition in the pseudoword condition (Shaywitz et al., 1998, 2002). The introductory sections of the three articles by Shaywitz et al. are exceptional in their clear statement of how their imaging studies assume and derive from cognitive models of word recognition. These studies adhere to the working hypothesis of cognitive neuroscience, as articulated by Raichle and coworkers and discussed above.

Of these studies, educators are most intrigued by the finding that a phonologically based reading intervention changes the functional anatomy of the brain in problem readers (Shaywitz et al., 2004). As a reading intervention, the Shaywitz et al. study adapted an instructional program designed to teach phonological skills that had previously been evaluated using reading-relevant performance (i.e., behavioral) measures (Blachman et al., 1999). The behavioral results reported in the Shaywitz imaging study replicate Blachman’s earlier behavioral study. The phonologically based program significantly improved reading fluency in disabled readers. In addition, the imaging study also found that after intervention, the reading-disabled children showed changes in brain activation patterns in areas previously associated with skilled reading (on the basis of previous lesion and imaging studies) both at the end of the intervention and one year later.

This is an impressive result, but what is the educational implication? The reading intervention itself is based solely on psychological and behavioral research. We also know from behavioral measures that the intervention improves reading in disabled readers. We know in fact that among 96 methodologically sound, published studies of systematic phonics instruction, the intervention used in the study by Shaywitz et al. ranks approximately tenth in mean overall effect size, fourth in word decoding effect size, and seventh in nonword reading effect size (Ehri et al., 2001). We also know based on a meta-analysis of these 96 studies that systematic phonics instruction has a moderate effect on reading outcomes, that effects are greater if instruction begins before first grade, that such instruction helps low- and middle-socioeconomic-status readers, and that it helped students at risk for reading disability (Ehri et al., 2001). We know that systematic phonics instruction works, that the program used in the imaging study is one of the stronger exemplars of such instruction, and that it results in development of neural circuits associated with skilled reading. Based on both the imaging study and previous behavioral research, our educational recommendation would most likely be that systematic phonics instruction should be implemented in beginning reading programs.

Suppose that the imaging study had shown no interpretable change in brain activations previously associated with skilled reading after intervention. How would our educational recommendation change? Given what we know about the educational impact of phonologically mediated reading instruction, based on behavioral measures of reading performance and the basic psychological research that supports this approach to reading instruction, our recommendation would change not at all. Given the vast body of experimental evidence that supports the importance of phonological awareness in reading, a “negative” imaging result would most likely be interpreted as showing that the imaging technology was not sufficiently sensitive to register the predicted postintervention changes in brain activation. The Shaywitz et al. result is of interest in as much as it is consistent with current cognitive models of reading and dyslexia—cognitive models the study assumes in its experimental design and tasks. But a negative imaging result in this case would not have implications for educational practice because the cognitive model is too well supported by psychological and behavioral studies.

For education, cognitive models matter more than identifying the brain areas that implement those models. Likewise, as Posner and Raichle argued (see above), using imaging studies to advance our understanding of the brain depends fundamentally on improving and refining our cognitive models. Cognitive neuroscience and imaging studies in particular, could provide insights that might help us refine our cognitive models. A few cognitive neuroscientists have started to discuss how brain activation patterns might provide constraints on our cognitive theorizing (Fiez and Petersen 1998; Fiez et al., 1999). However, imaging studies that go beyond establishing localization claims are complex, subtle, and quite rare within current cognitive neuroscience. They require an appreciation for the subtleties of competing cognitive models and the ability to interpret imaging results in the light of all relevant behavioral, neuropsychological, and imaging data available. Making such inferences goes well beyond correlating brain areas with cognitive functions and observing how activation patterns change after learning occurs. However, it is this more subtle and refined form of cognitive neuroscience that can build bridges from systems neuroscience to cognitive psychology. Once we can make this connection, we can explore possible new connections from the resulting cognitive models to the science of learning and educational practice.

Some will find the conclusions I draw about the implications of cognitive neuroscience and brain imaging for an improved science of learning discouraging. However, rather than enthusiastically speculating about the contributions that cognitive neuroscience and brain imaging are about to bring to instruction, we should be dismayed that cognitive models of reading and other subject domains, now nearly four decades old, have had such little impact in the classroom. Rather than suggest that schools of education are remiss in ignoring the neuroscience of reading, we should be concerned that educators remain unaware of the impact that cognitive psychological research has had and can have on educational practice.
**Cellular and Molecular Neuroscience: Action at a Distance**

The goal of education is learning. Learning is a memory process. Cellular and molecular neuroscientists are attempting to elucidate the molecular mechanisms underlying learning and memory by explaining memory processes at the synaptic level. In this reductive program, neuroscience is attempting to explain changes in behaviors that accure through learning in terms of changes in synaptic plasticity, or change. The leading candidate mechanisms for synaptic plasticity underlying learning and memory are long-term potentiation (LTP) and depression (LTD). LTP and LTD are the processes of stimulating a dendritic spine repeatedly, leaving it more or less responsive, respectively, to new input of the same type. What might be the educational implications of neuroscientific research at this level for teaching and learning?

Answering this question requires that we first answer another question: Does LTP cause observed changes in memory and behavior? Is it a causal mechanism for learning?

Typically in these discussions about possible causal mechanisms, neuroscientists attempt to analyze causal claims using necessity-sufficiency accounts of causal relations (Buonomano and Merzenich, 1998; Shors and Matzel, 1997).

So, one must establish that LTP is both a necessary and sufficient condition for the occurrence of long-term memory. Buonomano and Merzenich (1998) present criteria that must be satisfied to establish the existence of a causal link between concepts like LTP and changes in long-term memory.

LTP is a necessary condition for long-term memory, if, whenever there is a demonstrable change in behavior that can be attributed to the formation of a trace in long-term memory, LTP can be shown to occur in the appropriate neural circuit. LTP is a sufficient condition for memory formation if, whenever LTP occurs, there is a demonstrable change in behavior that can be attributed to the formation of a trace in long-term memory. For a somewhat different case, Buonomano and Merzenich (1998) detail the difficulties one confronts in attempting to establish empirically the existence of causal mechanisms at the synaptic level.

Shors and Matzel (1997), based on their review of the literature, concluded that LTP did not meet the criteria for providing a causal explanation of memory. To make a long argument very short, they documented instances where changes in memory occur without LTP and where LTP occurs without changes in memory. That is, they documented that LTP is neither necessary nor sufficient for memory change. Part of the problem resides in ambiguities over the definition of LTP. Another difficulty arises with experimental use of genetic and pharmacological interventions to block LTP that can have general, rather than specific, effects on the organism, making experimental interpretations difficult.

However, Shors and Matzel cite a more fundamental problem. They report that between 1974 and 1997, more than 1300 articles appeared that had “LTP” in their title. Of these, fewer than 80 described any behavioral manipulation relevant to assessing changes in memory. Furthermore, the articles that contained behavioral manipulations tended to provide evidence against the hypothesis that LTP is a memory mechanism. Thus, the claim that LTP is a molecular mechanism for learning and memory may be more of a dogma of neuroscientific memory research than a hypothesis that is being rigorously tested.

If so, cellular and molecular neuroscientific research on the causal mechanisms underlying memory and learning may represent a consistent set of claims under the dogmatic assumption, but in fact may be unconnected to memory phenomena as assessed by performance and behavioral data. If so, this undermines the relevance of the cellular and molecular results to learning and memory as studied by psychologists. And it is, after all, the psychological concept of memory that is most relevant to teaching and learning. If so, the implications of cellular and molecular neuroscience for teaching and learning are limited now, and will remain so, as long as neuroscientific research remains conceptually disconnected from psychology. Currently, LTP represents synaptic action at a considerable distance from memory and learning.

LTP is but a single example, but it is illustrative of a problem that arises whenever we attempt to link research across levels of analysis that range from molecules to behavior (Roediger et al., 2006). Are concepts like memory, learning, attention, retrieval, and plasticity as understood by neuroscientists the same concepts as understood by psychologists? Or are they only using the same words to designate different phenomena? If research at the synaptic level is to have implications for psychology and education, we need conceptual clarity at every interface between each level of analysis, from behavioral to molecular. Clearing away the semantic underbrush may be an important first step in outlining a research program wherein neuroscience, cognitive neuroscience, and psychology can have eventual implications for teaching and learning.

**CONCLUSION**

So I remain skeptical about the implications of neuroscience for education currently and into the near future. Maybe I should say the direct implications of neuroscience for education. I do believe that eventually we will be able to bridge neuroscience at its various levels of analysis with education, but I am convinced that all of these bridges will have at least one pier on the island of psychology.

**REFERENCES**


Points of View: On the Implications of Neuroscience Research for Science Teaching and Learning: Are There Any?^{1}

Anton E. Lawson

School of Life Sciences, Faculty of Organismal, Integrative and Systems Biology, Arizona State University, Tempe, AZ 85287-4501

NOTE FROM THE EDITOR

Points of View is a series designed to address issues faced by many people within the life sciences educational realm. We present differing points of view back to back on a given topic to stimulate thought and dialogue.

The focus of all contributed features and research articles in Issues in Neuroscience Education is the teaching and learning of neuroscience, from elementary school to graduate school audiences. However, neuroscience is unique as a branch of biology in that it includes the study of neuronal and brain mechanisms that may underlie learning. To highlight this unique position of neuroscience, we have chosen to focus this issue’s Points of View on how research findings in the field of neuroscience may or may not have implications for the teaching and learning of science in general. We invited authors to address the following questions:

• What are the current implications of neuroscience research, if any, for how to improve K–25 science teaching and learning in schools and universities?
• To what extent will neuroscience research into biological mechanisms of learning, memory, attention, and other brain functions inform educational practices and science teaching in the future?

INTRODUCTION

What, if anything, do teachers need to know about how the brain works to improve teaching and learning? After all, your plumber needs to know how to stop leaks—not the molecular structure of water. And we can learn how to use a computer without knowing how a computer chip works. Likewise, teachers need to know how to help students develop intellectually and learn—not necessarily how their brains work. Nevertheless, it is important for teachers to understand that what is being discovered about how brains work supports constructivist learning theory (Alexander and Murphy, 1999), which in turn supports inquiry-based teaching (American Association for the Advancement of Science, 1989; National Research Council, 1996, 2001; National Science Foundation, 1996). The goal of the present article is to explicate why this is so. Let’s start with some basics.

SOME BASICS OF BRAIN DEVELOPMENT

The neocortex, which is the most recently evolved part of the brain, has a full complement of brain cells (neurons) at birth—some 100 billion. Yet, the most rapid growth of the neocortex takes place during the first 10 years of life. This growth is primarily because of the proliferation of dendrites, i.e., the branching projections that connect with and receive input, via synapses, from nearby neurons. Importantly, the number of dendrites varies depending on use or disuse. For example, the neurons in the brain area that deals with word understanding (Wernicke’s area) have more dendrites in college-educated people than in people with only a high school education (Diamond, 1996). A classic study of the effect of disuse of neurons was conducted during the 1970s by Torsten Wiesel and David Hubel. They covered one of the eyes of newborn kittens at birth. When the covered eyes were uncovered 2 weeks later, they were unable to see. Presumably the lack of environmental input prevented the deprived neurons from developing normally. As Diamond (1996) put it, the phrase “use it or lose it” definitely applies in the case of neurons. Diamond adds, “No matter what form enrichment takes, it is the challenge to the nerve cells that is important. Data indicate that passive observation is not enough; one must interact with the environment.”

^{1} Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

DOI: 10.1187/cbe.06–03–0154
Address correspondence to: Anton E. Lawson (anton.lawson@asu.edu).
How does environmental interaction lead to an increase in the number of functional dendrites? According to neural network theory (Grossberg, 1982, 2005; Jani and Levine, 2000), dendrites become functional when neurotransmitter release rate increases at synaptic knobs. The increase in release rate makes signal transmission from one neuron to the next easier. Hence, learning is understood as an increase in the number of “operative” synaptic connections among neurons. That is, learning occurs when transmitter release rate at synaptic knobs increases so that the signals can be easily transmitted across synapses that were previously there, but inoperative. How, then, does experience strengthen connections?

HOW DOES EXPERIENCE STRENGTHEN CONNECTIONS?

Grossberg (1982, 2005) has proposed and tested equations describing the basic interaction of the key neural variables involved in learning. Of particular significance is his learning equation, which describes changes in transmitter release rate (i.e., \( Z_{ij} \)). The learning equation identifies factors that cause a change in \( Z_{ij} \) values. A forgetting or decay term, \( B_{ij} \), is a constant of decay. Thus, \( B_{ij} \) causes \( Z_{ij} \) to decay. The learning equation, which describes changes in transmitter release rate, is as follows:

\[
\dot{Z}_{ij} = -B_{ij}Z_{ij} + S'_{ij}[X_j]^+ 
\]

where the overdot represents a time derivative and \( i, j = 1, 2, \ldots, n \).

For example, consider Pavlov’s classical conditioning experiment in which a dog is stimulated to salivate by the sound of a bell. When Pavlov first rang the bell, the dog, as expected, did not salivate. However, upon repeated simultaneous presentation of food, which did initially cause salivation, and bell ringing, the ringing alone eventually caused salivation. Thus, the food is the unconditioned stimulus (US). Salivation upon presentation of the food is the unconditioned response (UCR). And the bell is the conditioned stimulus (CS). Pavlov’s experiment showed that when a CS (e.g., a bell) is repeatedly paired with a US (e.g., food), the CS alone will eventually evoke the UCR (e.g., salivation). How can the US do this?

Figure 1 shows a simple neural network capable of explaining classical conditioning. Although the network is depicted as just three cells (A, B, and C), each cell represents many neurons of the type A, B, and C. Initial food presentation causes cell C to fire. This creates a signal down its axon that, because of prior learning (i.e., a relatively large \( Z_{cb} \)), causes the signal to be transmitted to cell B. Thus, cell B fires, and the dog salivates. At the outset, bell-ringing causes cell A to fire and send signals toward cell B. However, when the signal reaches knob \( N_{AB} \), its synaptic strength \( Z_{AB} \) is not large enough to cause B to fire. So the dog does not salivate. However, when the bell and the food are paired, cell A learns to fire cell B according to Grossberg’s learning equation. Cell A firing results in a large \( S'_{AB} \) and the appearance of food results in a large \( E[X_B]^+ \). Thus, the product \( S'_{AB}E[X_B]^+ \) is sufficiently large to drive an increase in \( Z_{AB} \) to the point at which it alone causes node \( V_B \) to fire and evoke salivation. Food is no longer needed. The dog has learned to salivate at the ringing of a bell. The key theoretical point is that learning is driven by simultaneous activity of pre- and postsynaptic neurons, in this case activity of cells A and B.

ADAPTIVE RESONANCE: MATCHING INPUT WITH EXPECTATIONS

Another key aspect of neural network theory explains how the brain processes a continuous stream of sensory input by matching sensory input with expectations derived from prior experience. Grossberg’s mechanism for this, called adaptive resonance, is shown in Figure 2.

The process begins when sensory input \( X_{ij} \) is assimilated by a slab of neurons designated as \( F^{(j)} \). Because of prior experience, a pattern of activity, \( X_1 \), then plays at \( F^{(j)} \) and causes a firing of pattern \( X_2 \) at another slab of neurons \( F^{(j)} \). \( X_2 \) then excites a pattern \( X \) on \( F^{(j)} \). The pattern \( X \) is compared with the input following \( X_1 \). Thus, \( X \) is the expectation. \( X \) will be \( X_1 \) in a static visual scene and the pattern to follow \( X_1 \) in a temporal sequence. If the two patterns match, then you see what you expect to see. This allows an uninterrupted processing of input and a continued quenching of nonspecific arousal. Importantly, one is only aware of patterns that enter the matched/resonant state. Unless resonance occurs, coding in LTM is not likely to take place. This is because only in the resonant state is there both pre- and postsynaptic excitation of the cells at \( F^{(j)} \) (see Grossberg’s learning equation).

Now suppose the new input to \( F^{(j)} \) does not match the expected pattern \( X \) from \( F^{(j)} \). Mismatch occurs and this causes activity at \( F^{(j)} \) to be turned off by lateral inhibition, which in turn shuts off the inhibitory output to the nonspecific arousal source. This turns on nonspecific arousal and initiates an internal search for a new pattern at \( F^{(j)} \) that will match \( X_1 \).

Figure 1. Classical conditioning in a simple neural network. Cells A, B, and C represent layers of neurons.
Such a series of events explains how information is processed across time. The important point is that stimuli are considered familiar if a memory record of them exists at F(2) such that the pattern of excitation sent back to F(1) matches the incoming pattern. If they do not match, the incoming stimuli are unfamiliar and orienting arousal (OA) is turned on to allow an unconscious search for another pattern. If no such match is obtained, then no coding in LTM will take place unless attention is directed more closely at the object in question. Directing careful attention at the unfamiliar object may boost presynaptic activity to a high enough level to compensate for the relatively low postsynaptic activity and eventually allow a recording of the sensory input into a set of previously uncommitted cells.

HOW IS VISUAL INPUT PROCESSED IN DIFFERENT PARTS OF THE BRAIN?

As reviewed by Kosslyn and Koenig (1995), the ability to recognize objects visually requires participation of six major brain areas. As shown in Figure 3, sensory input from the eyes passes from the retina to the back of the brain and produces a pattern of electrical activity in the visual buffer (located in the occipital lobe). This activity produces a spatially organized image within the visual buffer. Next, a smaller region within the visual buffer (called the attention window) performs additional processing. The processed electrical activity is then simultaneously sent along two pathways on each side of the brain: two pathways run down (to the ventral subsystem in the lower temporal lobes), and two run up (to the dorsal subsystem in the parietal lobes). The ventral subsystem analyzes object properties, such as shape, color, and texture. The dorsal subsystem analyzes spatial properties, such as size and location. Patterns of electrical activity within the ventral and dorsal subsystems are then sent and matched to visual patterns stored in associative memory, which is located primarily in the hippocampus, the limbic thalamus, and the basal ganglia.
forebrain. If a good match is found (i.e., an adaptive resonance), the object is recognized and the observer knows the object’s name, categories to which it belongs, sounds it makes, and so on.

However, if a good match is not obtained, the object remains unrecognized and additional sensory input must be obtained. Importantly, the search for additional input is not random. Rather, stored patterns are used to make a second hypothesis about what is being observed, and this hypothesis leads to new observations and to further encoding. In the words of Kosslyn and Koenig, when additional input is sought, “One actively seeks new information that will bear on the hypothesis. The first step in this process is to look up relevant information in associative memory” (p. 57). This information search involves activity in the prefrontal lobes in an area referred to as working memory. Activating working memory causes an attention shift of the eyes to a location where an informative component should be located. Once attention is shifted, the new visual input is processed in turn. The new input is then matched to shape and spatial patterns stored in the ventral and dorsal subsystems and kept active in working memory. Again, in Kosslyn and Koenig’s words, “The matching shape and spatial properties may in fact correspond to the hypothesized part. If so, enough information may have accumulated in associative memory to identify the object. If not, this cycle is repeated until enough information has been gathered to identify the object or to reject the first hypothesis, formulate a new one, and test it” (p. 58).

For example, suppose while driving your car you observe what seems to be a puddle of water in the road ahead. Thanks to connections in associative memory, you know that water is wet. So when you continue driving, you expect that your tires will splash through the puddle and get wet. But upon reaching the puddle, it disappears and your tires stay dry. Therefore, your brain rejects the puddle hypothesis and generates another hypothesis, perhaps a mirage hypothesis. The pattern of information processing involved in this example can be summarized as follows:

If . . . the object is a puddle of water,
and . . . you continue driving toward it,
then . . . your tires should splash through the puddle and they should get wet.
But . . . upon reaching the puddle, it disappears and your tires do not get wet.
Therefore . . . the hypothesis is not supported; the object was probably not a puddle of water.

In other words, as one seeks to identify objects, the brain generates and tests stored patterns selected from memory. Kosslyn and Koenig even speak of these stored patterns as hypotheses, where the term hypothesis is used in its broadest sense. Thus, brain activity during visual processing uses an If/then/Therefore hypothetico-deductive pattern. One looks at part of an unknown object and the brain spontaneously and immediately generates an idea of what it is—a hypothesis. Thanks to links in associative memory, the hypothesis carries implied consequences (i.e., expectations/predictions). Consequently, to test the hypothesis one can carry out a simple behavior to see whether the prediction does in fact follow. If it does, one has support for the hypothesis. If it does not, then the hypothesis is not supported and the cycle repeats.

**IS AUDITORY INPUT PROCESSED IN THE SAME HYPOTHETICO-DEDUCTIVE WAY?**

The visual system is only one of several of the brain’s information processing systems. However, information seems to be processed in a similar hypothetico-deductive manner by other brain systems. For example, with respect to learning the meaning of spoken words, Kosslyn and Koenig (1995) state “Similar computational analyses can be performed for visual object identification and spoken word identification, which will lead us to infer analogous sets of processing subsystems” (p. 213).

Details of this hypothesized word recognition subsystem are not important. Rather, what is important is that word recognition, like visual recognition, involves brain activity in which hypotheses arise immediately, unconsciously, and before any other activity. In other words, the brain does not make several observations before it generates a hypothesis of what it thinks is out there. Instead, from the slimmest piece of input, the brain immediately generates an idea of what it “thinks” is out there. The brain then acts on that initial idea until subsequent behavior is contradicted. In other words, the brain is not an inductivist organ. Rather, it is an idea-generating and -testing organ that works in a hypothetico-deductive way. There is good reason in terms of human evolution why this would be so. If you were a primitive person and you look into the brush and see stripes, it would certainly be advantageous to get out of there quickly as the consequences of being attacked by a tiger are dire. And anyone programmed to look, look again, and look still again in an “inductivist” way before generating the tiger hypothesis would most likely not survive long enough to pass on his plodding inductivist genes to the next generation.

The important point is that learning does not happen the way you might think. Your brain does not prompt you to look, look again, and look still again until you somehow internalize a successful behavior from the environment. Rather, your brain directs you to look and, as a consequence of that initial look, the brain generates an initial hypothesis that then drives behavior, behavior that carries with it a specific expectation. Hopefully, the behavior is successful in the sense that the prediction is matched by the outcome of the behavior. But sometimes it is not. So the contradicted behavior then prompts the brain to generate another hypothesis and so on until eventually the resulting behavior is not contradicted. In short, we learn from our mistakes—from what some would call trial and error.

**CAN NEURAL NETWORKS EXPLAIN HIGHER LEVELS OF REASONING AND LEARNING?**

Research has shown that the previous neural network principles can be successfully applied to explain more complex learning. For example, Levine and Prueitt (1989) developed and tested a neural network model to explain performance of normal persons and those with frontal lobe damage on...

In 1610 in his Sidereal Messenger, Galileo reported observations made by a new telescope of his invention. In the report Galileo claims to have discovered four “planets” circling Jupiter. As he put it: “I should disclose and publish to the world the occasion of discovering and observing four planets, never seen from the beginning of the world up to our times” (Galilei, 1610, as translated and reprinted in Shapley et al., 1954, p. 59).

Unlike most modern scientific papers, Galileo’s report is striking in the way in which it chronologically reveals the steps in his thinking. Thus, it provides an extraordinary opportunity to gain insight into the thinking involved in an important scientific discovery. What follows is a brief recapitulation of part of that report followed by an attempt to fill in gaps in Galileo’s reasoning as he interpreted his observations. Galileo’s reasoning will then be modeled in terms of Kosslyn and Koenig’s neural network principles. Let’s start with Galileo’s initial observations on January 7.

January 7

Galileo made a new observation on January 7 that he deemed worthy of mention. In his words, “I noticed a circumstance which I had never been able to notice before, owing to want of power in my other telescope, namely that three little stars, small but very bright, were near the planet (i.e., Jupiter).”

This statement suggests that Galileo’s observation was immediately assimilated by a fixed star category. In other words, he knew from past experiences that some of the objects in the night sky were fixed stars (i.e., stars that were part of the unchanging celestial sphere). But Galileo’s continued thinking led to some initial doubt as this following remark reveals: “. . . and although I believed them to belong to the number of the fixed stars, yet they made me somewhat wonder, because they seemed to be arranged exactly in a straight line, parallel to the ecliptic, and to be brighter than the rest of the stars, equal to them in magnitude.”

Why would this observation lead Galileo to somewhat wonder? Perhaps he was reasoning along these lines:

If . . . the three objects are fixed stars, and . . . their sizes, brightness, and positions are compared with each other and to other nearby stars, then . . . variations in size, brightness, and position should be random, as is the case for other fixed stars.

But . . . “they seem to be arranged exactly in a straight line, parallel to the ecliptic, and to be brighter than the rest of the stars.”

Therefore . . . the fixed-star hypothesis is not supported. Or as Galileo put it, “yet they made me wonder somewhat.”

January 8

The next night Galileo made another observation. Again, in his words: “. . . when on January 8, I found a very different state of things, for there were three little stars all west of Jupiter, and nearer together than on the previous night, and they were separated from one another by equal intervals, as the accompanying figure shows.”

(East) ○ * * * (West)

The new observation puzzled Galileo and raised another question. Again, in Galileo’s words: “At this point, although I had not turned my thoughts at all upon the approximation of the stars to one another, yet my surprise began to be excited, how Jupiter could one day be found to the east of all the aforementioned stars when the day before it had been west of two of them.” Presumably this observation was puzzling because it was not the expected one based on his fixed-star hypothesis.

Galileo continues, “. . . forthwith I became afraid lest the planet might have moved differently from the calculation of astronomers, and so had passed those stars by its own proper motion.” This statement suggests that Galileo has not yet rejected the fixed-star hypothesis. Instead, he has generated an ad hoc hypothesis that the astronomers made a mistake, i.e., perhaps their records were wrong about how Jupiter moves relative to the fixed stars in the area. This hypothesis could subsequently be tested as follows:

If . . . the astronomers made a mistake, and . . . I observe the next night, then . . . Jupiter should continue to move east relative to the stars, and the objects should look like this: (East) ○ * * * (West)

Of course, we cannot know whether this is what Galileo was thinking, but if he were thinking along these lines, he would have had a very clear prediction to compare with the observations he hoped to make the next night.

January 9 and 10

Galileo continues: “I therefore waited for the next night with the most intense longing, but I was disappointed of my hope, for the sky was covered with clouds in every direction. But on January 10th the stars appeared in the following position with regard to Jupiter, the third, as I thought, being hidden by the planet.”

(East) * * ○ (West)

What conclusion can be drawn from this observation in terms of the astronomers-made-a-mistake hypothesis? Consider the following reasoning:

If . . . the astronomers made a mistake, and . . . I observe the next night, then . . . Jupiter should continue to move east relative to the “stars,” and the objects should look like this:

(East) ○ * * * (West) (expected result)

But . . . the objects did not look like this, instead they looked like this:

(East) * * ○ (West) (observed result)

Therefore . . . the astronomers-made-a-mistake hypothesis is not supported.
Interestingly, Galileo states:

When I had seen these phenomena, as I knew that corresponding changes of position could not by any means belong to Jupiter, and as, moreover, I perceived that the stars which I saw had always been the same, for there were no others either in front or behind, within the great distance, along the Zodiac—at length, changing from doubt into surprise, I discovered that the interchange of position which I saw belonged not to Jupiter, but to the stars to which my attention had been drawn.

(p. 60)

So, Galileo concluded that the astronomers had not made a mistake, i.e., the changes of position were not the result of Jupiter’s motion. Instead, they were due to motions of the “stars.”

January 11 and Later

Galileo is now left with the task of formulating and testing another hypothesis. The following observation and remarks make it clear that he did not take long to do so:

Accordingly, on January 11 I saw an arrangement of the following kind:

(East) * * ○ (West)

namely, only two stars to the east of Jupiter, the nearer of which was distant from Jupiter three times as far as from the star to the east; and the star furthest to the east was nearly twice as large as the other one; whereas on the previous night they had appeared nearly of equal magnitude. I, therefore, concluded, and decided unhesitatingly, that there are three stars in the heavens moving about Jupiter, as Venus and Mercury round the sun.

(p. 60)

Galileo’s remarks make it is clear that he has “conceptualized” a situation in which these objects are traveling around Jupiter in a way analogous to the way our moon travels around the Earth. Thus, he has rejected the fixed star hypothesis and accepted an alternative hypothesis in which the objects are traveling around Jupiter—they are moons of Jupiter. How could Galileo have arrived at such a conclusion? Consider the following reasoning:

If . . . the objects are orbiting Jupiter, and . . . I observe the objects over several nights, then . . . some nights they should appear to the east of Jupiter and some nights they should appear to the west. Further, they should always appear along a straight line on either side of Jupiter.

And . . . this is precisely how they appeared.

Therefore . . . the moons-of-Jupiter hypothesis is supported.

Galileo’s previous statement continues as follows:

. . . which at length was established as clear as daylight by numerous other subsequent observations. These observations also established that there are not only three, but four, erratic sidereal bodies performing their revolutions round Jupiter . . . These are my observations upon the four Medicean planets, recently discovered for the first time by me.

(pp. 60–61)

MODELING GALILEO’S REASONING

Kosslyn and Koenig’s description of brain subsystem functioning is about recognizing objects present in the visual field during a very brief time period—not distant spots of light seen through a telescope. Nevertheless, the hypo-

Figure 4. What might have been in Galileo’s working memory when he tested the moons hypothesis?
thetico-deductive nature of this system functioning is clear. And all one need do to apply the same principles to Galileo’s case is to extend the time frame over which observations are made—observations that either match or mismatch expectations. For example, Figure 4 shows how the brain subsystems may have been involved in Galileo’s reasoning as he tests his moons hypothesis.

The figure highlights the contents of Galileo’s working memory, which is seated in the lateral prefrontal cortex, in terms of one cycle of If/then/Therefore reasoning. As shown, to use If/then/Therefore reasoning to generate and test his moon hypothesis, Galileo must not only allocate attention to it and its predicted consequences, he must also inhibit his previously generated fixed-stars and astronomers-made-a-mistake hypotheses. Thus, working memory can be thought of as a temporary network to sustain information while it is processed. During reasoning, one must pay attention to task-relevant information and inhibit task-irrelevant information. Consequently, working memory involves more than simply allocating attention and temporarily keeping track of it. Rather, during the reasoning process, working memory actively selects information relevant to one’s goals and actively inhibits irrelevant information.

INSTRUCTIONAL IMPLICATIONS

How then do people learn? The answer seems to be through encountering puzzling observations and trying to explain them through cycles of If/then/Therefore hypothetico-deductive reasoning. Presumably, this is because this is the way that the brain spontaneously processes information. The more skilled people are at reasoning in this manner, the better they are at learning, at constructing new knowledge (Lawson et al., 2000). The key point in terms of instruction is that for meaningful and lasting learning to occur, students must personally and repeatedly engage in the generation and test of their own self-generated ideas. This means that laboratory and field-based activities become the main instructional vehicles. But such activities cannot be “cookbook” in nature. Instead, they should allow students the freedom to openly inquire and raise puzzling observations. The puzzling observations should then prompt students to generate and test their own alternative explanations with the following sorts of questions becoming the central focus of instruction:

- What did you observe?
- What is puzzling about what you observed?
- What questions are raised?
- What are some possible answers/explanations?
- How could these possibilities (alternative hypotheses) be tested?
- What does each hypothesis and planned test lead you to expect to find?
- What are your results?
- How do your results compare with your predictions?
- What conclusions, if any, can be drawn?

ACKNOWLEDGMENTS

This material is supported in part by the National Science Foundation under award EHR 0412537.

REFERENCES


The Faculty Institutes for Reforming Science Teaching (FIRST II) (http://www.first2.org) network was established in 1998 by ecologists accustomed to conducting their research in the field. The original impetus behind establishing FIRST was to engage biology faculty in active, inquiry-based science teaching, first in the field, an ideal environment for cooperative learning, and then in their classes. The impact of FIRST is seen in the activities of faculty now from all areas of biology who are change agents for the improvement of undergraduate science education, but major challenges toward implementing change still loom. Like many proponents of change, we are flush with new course and curricular materials and active-learning instructional strategies, but short on substantive research to back our claims that the “new” teaching approaches promote student learning better than traditional approaches. Many faculty remain unconvinced of the need for change because they have not been confronted with enough assessment data demonstrating that the new instructional methods have a positive impact on student learning—it is the “show me the data” dilemma. In fact, many faculty do have substantive assessment data, and others would like to collect such data. The challenge is learning what kinds of questions to ask; how to gather and analyze assessment data; and, ultimately, how to disseminate the results to an appropriate target audience.

One way to meet these challenges has been through publication of assessment results in a peer-reviewed educational journal such as CBE—Life Sciences Education, whose registered subscriber base is growing at a rapid rate. However, it is difficult for such a journal to provide sufficient guidance in assessment strategies and techniques to potential authors new to educational research. We present the case here for a somewhat different approach toward meeting these challenges: the creation of an assessment database in biology education.

The fundamental principles for the FIRST database are modeled after those described for the Long-Term Ecological Research Network (http://lternet.lternet.edu/DTD/) database and GenBank (http://www.ncbi.nlm.nih.gov/Genbank/index.html). To be successful, such a database must contain data desired by and useful to a large group of users, data must be easily added, the level of technological sophistication needed to use the system must match that of the users, and there must be a mechanism for submitting and dealing with queries that identify user needs (Michener et al., 1997).

Open source programs associated with the FIRST database will provide faculty with easy, efficient ways to document and code their data for uploading to the database. Faculty can upload student response data and download coded student data in a variety of file formats designed to work with programs such as Excel, Access, Filemaker Pro, SAS, and SPSS. As faculty gain experience using the database, they can extract data from different courses and institutions to conduct comparative analyses. For example, faculty can search and download questions and student responses from assessments on evolution that include multiple-choice instruments and extended response questions. Investigators may request restricted access to their data for a limited time of up to 3 yr, to encourage use of the database while giving them time to publish. After the release date, data will be available to all other database users. Documentation for all data retrieved from the database will contain source metadata including the names, e-mails, and institutions of the submitted data and links to any publications associated with the data.

DOI: 10.1187/cbe.05–12–0133
Address correspondence to: Diane Ebert-May (ebertmay@msu.edu).
The database accepts a variety of assessment types, including multiple choice; "clicker data"; short answer; original text of essays; coded responses of essays; computer-generated files, including scanned student drawings; and concept maps. The database is dynamic, which means that as new assessment formats are developed, new schema files (similar to Endnote connection files) can be generated, ensuring future database compatibility. Before data are uploaded, information that could identify an individual student will be replaced with a random number to "de-identify" data in compliance with Institutional Review Board regulations. Only faculty will maintain the ability to decode student identifiers locally.

Currently, we are field-testing the database by uploading and downloading multiple types of assessments and students’ responses. For example, we have added the Concept Inventory of Natural Selection (Anderson et al., 2002) and assessments for the carbon cycle (Ebert-May et al., 2003) to the database, with metadata describing questions, foils (for multiple choice), and extended responses. The metadata show the specific content area for each question, for example, 1) populations evolve, not individuals; and 2) genetic traits are inherited, not acquired. Eventually, investigators may download files and tables of student responses that are classified according to these concepts and then import formatted data into Excel or other software for analysis.

Faculty can also mine raw student responses from many courses and institutions and recode the data to meet their research objective. As a hypothetical example, suppose an instructor predicts that understanding DNA replication and protein synthesis is vital to learning evolution by natural selection. To test the hypothesis, she needs to match assessments of DNA replication and protein synthesis to assessments of evolution. First, the instructor queries the database and finds all courses that include assessments for both concept areas; 50 courses from 20 institutions are available. After examining the course and instrument metadata, the instructor narrows the data set further to majors’ courses that used pre- and posttests, thereby reducing the data set to 30 courses from 10 institutions. When the instructor downloads the data for analysis, the query is saved with an accession number. This number acts as a record that can be used when double-checking results and can be cited in published articles.

The FIRST database provides instruments and protocols for conducting classroom research, as well as data for cross-institutional studies of student learning and faculty teaching. However, the question remains, "If we build it, will they come?" We are currently studying the nature, extent, and effects of faculty change among a large number of faculty who participated in two different professional development programs. If faculty do change to improve their teaching practice based on learning theory and tested pedagogical principles, will student learning in fact improve? The FIRST assessment database will help to explore that question. It will also provide a useful source of assessment methods, data, and analyses to faculty who want to improve their teaching and measure the results, or simply to convince their colleagues that improving teaching can significantly affect student learning. The prototype for the database is in place; if the project is funded, a public version will be available within 12 mo. In the meantime, a subsample of faculty from the American Society for Cell Biology were surveyed about their interest in using the database. We welcome any feedback and will solicit input from individuals about the database, so together we can improve undergraduate life science education.

ACKNOWLEDGMENTS

We thank the National Science Foundation for its long-term support of the FIRST project, Faculty Institutes for Reforming Science Teaching (DUE 0088847).

REFERENCES


Essay

A Simple E-Mail Mechanism To Enhance Reflection, Independence, and Communication in Young Researchers

A. Malcolm Campbell*† and Barbara Lom*‡

*Biology Department, †Genomics Program, and ‡Neuroscience Program, Davidson College, Davidson, NC 28035

Submitted June 21, 2006; Revised July 31, 2006; Accepted August 5, 2006
Monitoring Editor: Mary Ledbetter

Providing undergraduates with mentored research experiences is a critical component of contemporary undergraduate science education. Although the benefits of undergraduate research experiences are apparent, the methods for mentoring young scientists as they first begin navigating the research lab environment are reinvented in labs all over the world. Students come to research labs with varied skills, motivations, needs, and dispositions, placing each student and mentor in a unique relationship. How can we help students become aware of their own intellectual progress? How can we encourage our students to take initial steps toward independent investigation? When do we need to let setbacks happen? We have developed a simple mechanism to address these common problems. Each week, students in our labs answer a series of five questions by e-mail that improve lab communication and help students develop into mature scientists without taxing an instructor’s already busy schedule. Our observations, experiences, and student feedback indicate that this approach is a useful mechanism to help faculty who mentor young scientists in the research lab.

INTRODUCTION

Many successful scientists trace the origins of their research careers back to influential experiences as undergraduates working in their first research laboratory. This formative time in a laboratory, when students first do original research, can shape the futures of budding scientists in powerful ways. Current science pedagogy strongly recommends integrating research into the undergraduate curriculum both as a means to educate the public in the methods of scientific inquiry and as a way to stimulate the next generation of scientists (Tobias, 1992; National Research Council, 2003). Moreover, undergraduate or postbaccalaureate research experience is expected for admission to most graduate programs in the sciences. Consequently, providing undergraduates effective mentoring is a critical component of contemporary undergraduate science education (Pfund et al., 2006). Although the benefits of undergraduate research experiences are apparent, the methods for mentoring scientists as they first begin navigating the research lab environment are far less clear. Few mentors received training in effective ways to mentor young students who need to learn not only a lab’s specific techniques but also the intellectual methods of approaching research problems and the culture of the research lab environment.

Considerable attention has been focused on measuring the impact of student research on preparing future scientists (Tobias, 1992; National Research Council, 2003; Lopatto, 2004), yet it remains difficult to determine how good mentoring in laboratory training affects individual students. What can we do, as scientist-educators, to improve our students’ research experiences? How can we help students become more aware of their own intellectual progress, which may foster higher levels of learning (Bloom, 1956; Magolda, 1992, 2001; Metcalfe and Shimamura, 1994; King and Kitchener, 1994; Kronholm, 1994)? How can we encourage our students to take initial steps toward independent investigation? If we knew what our students were thinking, would we be better at letting them learn by their own mistakes when minor setbacks happen?

Few of us who supervise new researchers have formal training in mentoring; consequently, most mentors develop mentoring styles by drawing from their own experiences as protégées (Plund et al., 2006). It is a challenge to find an

DOI: 10.1187/cbe.06–06–0170
Address correspondence to: A. Malcolm Campbell (macampbell@davidson.edu).
appropriate balance between micromanaging our students and providing them with independence and space to learn how to be successful in the lab. In addition, students come to research labs with varied skills, motivations, needs, and dispositions, placing each student and mentor in a unique relationship.

REFLECTIVE QUESTIONS IN A RESEARCH LAB

To enhance communication, comprehension, reflection, and independence among our undergraduate research students, we have developed a simple mechanism for gauging student growth, enhancing communication, and modeling the higher levels of thinking necessary for successful research. On a regular basis, we ask our research students to answer these short questions via e-mail:

1. How have you spent your time?
2. What do you know?
3. What don’t you know?
4. How can you find out what you don’t know?
5. What are your frustrations?

The first two questions stimulate students to document their progress, questions three and four encourage students to identify gaps in their knowledge and ways to fill in those gaps, and the last question allows students to identify and share any roadblocks they encounter in their research and learning.

How Have You Spent Your Time?

By summarizing how their lab time was spent (some of which may occur out of the mentor’s sight), the mentor can evaluate student participation and efficiency. Although this documentation is not meant to be a time sheet, it can be very useful in assigning participation grades or noting trends when the level of student engagement may be changing and needs to be addressed in some way. Some new students in the lab are very good at streamlining their work habits, reaching goals, and producing data, whereas others can get sidetracked, overwhelmed, or lost. This straightforward documentation of how lab time is spent can also help the mentor and student identify any issues with efficiency, time management, and research priorities that may arise. Having a student-generated record of the time spent on particular tasks can help the mentor identify areas for improvement. Moreover, question number one can help students appreciate the progress in their skills. For example, a student may need all day to complete a procedure when he or she first learns it, but with time and experience the new investigator may be able to accomplish the same procedure much more rapidly and accurately. By regularly describing how they spend their lab time, students can see how they are improving their technical competence, an important and encouraging accomplishment in the early stages of research that often precedes the production of useful data.

What Do You Know?

By stating what they have learned recently, students can document and appreciate how much they have learned via their research experiences. Student responses to this question range widely from acquisition of specific facts and familiarity with relevant research literature to proficiency in technical skills, protocols, and experimental design. Sometimes they comment on common novice pitfalls, specific insights, presentation strategies, or how to interact with others in the lab. Because the excitement of research findings can be sparse, particularly for new researchers, students often lose sight of their own progress. We find that this question sets an upbeat tone and helps them document new lessons learned since their last entry. In our experience, students have little trouble identifying new knowledge. Moreover, this regular statement of what students have learned in the past week helps both students and mentors appreciate the intellectual gains students are making as they go through the research process.

What Don’t You Know?

It is not unusual for students to feel anxious in a new lab environment, but many have trouble identifying the cause of their apprehension. One source of this discomfort can be an inability to recognize holes in their knowledge. Identifying what they do not know is an important first step for students learning to take charge of their own education, think independently, and develop problem-solving strategies. In our experience, college students are so accustomed to being tested on what they know (or are supposed to know), that sometimes they have trouble adjusting to this situation where they are expected to identify the gaps in their own knowledge. We have found that many students come to appreciate this nonthreatening and unusual opportunity to recognize what they do not know. By explicitly encouraging students to define the specific gaps in their knowledge, we can help students acknowledge and approach their uncertainty in a way that encourages them to communicate, problem solve, and ultimately become more productive scientists.

From the mentor’s perspective, this question of what a student does not know often identifies critical areas where the mentor may have inadvertently assumed knowledge that the students do not yet have, where students misunderstood important information, or where expectations may have been unclear. We have found this question to be particularly important in our mentoring relationships because many students are more comfortable revealing ignorance via private written communication than speaking in public or even one-on-one. We can respond to their issues directly by supplying answers, or by directing students to appropriate references and resources. In some instances, the entire lab group needs to be informed, so the gaps students identify in this question can be used to structure lab meetings (or other shared time) without putting a student on the spot. Regardless of the format of our response, we can address student needs quickly, thereby reinforcing the value of their communicating openly with their supervisors. Question four (How can you find out what you don’t know?) builds on question three by encouraging students to consider active strategies for filling in gaps. Knowing how to find an answer to an open question is an important skill that all successful researchers need to develop.
How Can You Find Out What You Don’t Know?

Original research requires technical skills and intellectual maturity in order to trouble shoot, interpret data, and make progress. Few students develop all these traits via traditional lecture and lab courses. In the research lab, the sink-or-swim approach is a common mechanism by which new research students acquire these important skills. Many successful investigators are living testament that a sink-or-swim approach can identify future scientists. Although some students may be natural swimmers who thrive in the research lab, there are students who may not be able to swim on their own at first, but with guidance can be taught to swim, and eventually thrive in the research lab environment. Many people and organizations have called for increased diversity in the science workforce to improve the talent pool of tomorrow’s scientists (National Research Council, 2003; Lawrence, 2006). If we continue to apply the same sink-or-swim selection pressure to all students, can we realistically expect different outcomes of our educational system and thus a more diverse population of scientists? It seems clear that we need to provide a range of mentoring options if we are to respond appropriately to the reality that our students are young adults with diverse backgrounds, personalities, and needs.

Rather than fostering recurrent bad habits in weak students, we see questions three and four as a way to support students who need opportunities to grow without hindering the students who may not need such explicit opportunities. By devising their own solutions to self-defined problems, our students begin to take charge of their own education, gaining confidence and independence, and learning perhaps the most important of all skills in science—problem solving. By identifying areas of uncertainty and devising ways to address them, students can raise their level of self-awareness and mature as independent and life-long learners.

What Are Your Frustrations?

Unlike the previous four questions, this question opens the door for students and mentors to address personal problems related to lab research. Many problems that begin as failures of interpersonal communication can result in situations that reduce the efficiency with which the laboratory operates (Cohen and Cohen, 2005). In addition to personality conflicts, students often reveal frustrations over time constraints, poor planning, broken equipment, missing reagents, or inconsiderate labmates. It seems odd that students are reluctant to share such concerns in person, but we have learned from our experiences that some students have difficulty expressing themselves verbally or fear that it is impolite to reveal a weakness or “tattle” on a labmate with whom they are having difficulties. Therefore, the written answers to this question can uncover behavior patterns or communication breakdowns that the students are experiencing but may often be hidden from a busy mentor’s view.

Problems between labmates can be addressed early if intervention is required. However, the more common outcome is more effective individualized mentoring. Each student has different needs, and a one-size-fits-all approach to personnel management is often inadequate, especially for younger students. In addition, we want to provide our students with a role model for supervising people and projects by supporting those who would have sunk if left alone. We hope we are encouraging some students to consider a life of research instead of succumbing to pressures to pursue high-profile careers such as medicine, business, or law. For those students who do not continue in science or research, we expect that these five questions will also be a valuable strategy they can use to solve problems, reflect on their progress, and communicate with coworkers in other disciplines as well.

LOGISTICS OF QUESTIONS

Principal investigators typically are busy people with responsibilities that extend far beyond supervising new students in their labs. Consequently, our five-question approach may sound like more busywork that will add to e-mail accounts that are already overwhelming. However, the time commitment of our approach is minimal and the payoff substantial, even time saving, for student learning and meeting our research goals.

We have found that a once-a-week e-mail works well for independent study or group investigation research courses during the academic year. For full-time summer research students, daily answers combined into one week-long document submitted Friday afternoons works very well (Figure 1). Because the submissions are by e-mail, it is easy for us to reply to minor issues. For those issues that require more direct or in-depth communication, we can set up appointments and/or use the next scheduled lab meeting as a forum to reach everyone without singling out an individual student. An added benefit of e-mail correspondence is that we can archive these e-mails in a folder and refer back to them as necessary for documenting student progress for use in letters of recommendation, considering class participation grades, and personal reflections of a semester.

We do not grade student answers to these questions because we view this communication as a mechanism for mentoring rather than evaluating. In fact, we do not require complete sentences or perfect grammar. The questions are meant to foster independence, reflection, and open commu-
communication. We feel that answers submitted without formal judgment are more likely to foster honest communication because students often feel that they need to “perform” by supplying the “right answers” for graded assignments. We use these five questions simply and directly to foster trusting and open relationships with our research students and enhance the research productivity of our labs. We hope this approach will lead to stronger relationships and mutual respect that can grow into long-lasting professional relationships.

STUDENT FEEDBACK

This “five questions” method was piloted during the summer of 2005 with four research students. During the 2005–2006 academic year, we used these questions in two group investigation courses. The courses encouraged students to learn new research techniques beyond the scope of traditional lab courses (designing and printing DNA microarrays, immunostaining, confocal microscopy, tissue culture, etc.) in order to answer novel research questions. Students arranged independent work times throughout the week and met in groups with their respective instructor in a weekly lab meeting format. We used the e-mail to identify issues that needed to be addressed during the weekly lab meeting. The five-question e-mails revealed many important issues that could be addressed easily during the lab meeting such as allocation of research time, clarifying research objectives, assigning research tasks, scheduling training times, and addressing conceptual questions. Any lingering or individual questions were addressed by e-mail or in person. At the end of the semester, students commented favorably on anonymous evaluation sheets that asked if the weekly e-mail assignment helped them reflect on what they were learning and communicate with the instructor. Specific comments about the e-mail assignment included:

- “They keep us on track.”
- “I found the five questions an excellent time for reflection, reevaluations, and planning for the upcoming week. They also provided a low-stress way to express concerns.”
- “Lab meetings were very productive. We got done what we needed to get done quickly and efficiently.”
- “I thought the five questions were great, especially the ones that asked us what we couldn’t do and how we could go about learning how to do them. This encouraged us to think for ourselves.”
- “I think this was very helpful for me. First of all, it made me actually think about what I had learned/not learned and to put my frustrations into words. It was also helpful to let you know what was going on in the lab if we were having problems.”

Other comments on the end-of-the-semester evaluations indicated that many of our learning goals of teaching independent thinking and problem solving were achieved, even though the evaluation form did not ask about these goals explicitly. For example, students commented:

- “This whole class involved thinking for ourselves.”
- “Communication was key to successfully completing projects in this course.”
- “I also have a greater appreciation for the importance of teamwork and accountability in the lab. Honesty and reliability are so key to a successful research group.”
- “Constant communication was key. . . . I think if I had been trying to go through all of these experiments and processes by myself, I would have been way more frustrated and way less successful. . . . ”
- “I’ve realized that I really do love researching in the lab and that I can handle the frustrations.”

Finally, student responses to a question on advice they would give on being successful in a research course to a friend taking the course in the future revealed that communication, planning, and identifying personal strengths were important lessons learned by our students.

- “I would say that it is important to learn how to communicate with your peers early on. This will be invaluable in the future. Also, plan ahead. If you have your project organized, then you can set goals for yourself, as well as deadlines. That way, when the deadline approaches, you are not frantically rushing to get everything done.”
- “Ask questions. Plan ahead. Always leave more than enough time to complete the project at hand. Start working early! Communicate daily with lab mates about progress, questions, tips. Establish good relationships with all lab mates.”
- “Start communication with the group members as soon as possible. It helps so much to use each other’s strengths . . . . to get the tasks completed. Don’t be afraid to ask for help. The sooner you know if you’re going in the right direction, the better things will be.”

CONCLUSIONS

Asking students to answer these questions is not intended to replace good lab management or dedicated mentoring. We found this efficient, five-question e-mail mechanism goes a long way to establishing good working relationships, open communication, and demonstrating the value of regular reflection on one’s progress and challenges with our students. Furthermore, these questions have improved our ability to gauge student progress and attitudes toward research by providing an important window into the minds of our students. When our primary goal is helping students learn more effectively, teaching them how to take the lead in their own education is a beneficial outcome (that does not often appear on standardized tests or course evaluations). The amount of time it takes to implement these questions is directly proportional to the size of the lab. Mentors with more students must spend more time reading and responding to questions, but that time would be well spent if the mentor becomes more fully aware of each student’s progress. Regardless of lab size, we feel these questions are a very effective means to improve lab communication and efficiency for young research students. Moreover, these questions have the important potential to reach those students who, due to differences in personalities, are reluctant to contribute during lab meetings or interrupt a busy mentor. If a student knows he or she will be heard at least once a week, the mentor may see increased confidence and public
communication. Most importantly, these questions help foster a learning environment that is rewarding, enjoyable, and stimulating for individual growth—exactly the type of environment in which students can discover the joys of science.

Our easy-to-use method is consistent with efforts to use reflective thinking to enhance student learning. For example, Blank (2000) found improved metacognition and the duration of student learning when teachers provided students with additional discussions about their education. Blank (2000) found no difference in the amount of knowledge content as a result of student-teacher discussions on learning. However, using e-mail as a less time-intensive mechanism of communication can improve the quantity of student learning (Smith et al., 1999; Yu and Yu, 2002). Interestingly, Yu and Yu (2002) measured an improvement in knowledge content as a result of additional e-mail communication, but student attitudes toward the subject were not improved. These different results highlight the difficulty in assessing educational gains that can deter many educators from adopting new methods. However, each of these studies demonstrated learning improvements of some kind. Furthermore, Pfund et al. (2006) demonstrated that supervisor communication with student researchers is critical to improving the mentoring relationship. Mentors were encouraged to help students gain confidence and independence. In short, regularly scheduled queries of students about their own learning are simple to implement and can improve a student’s research experience. Therefore, we will continue to use this reflective e-mail method as one tool to improve undergraduate research training/mentoring.

ACKNOWLEDGMENTS
We thank all the students who have worked in our labs and Drs. Laurie Heyer and Christine Broussard for critical discussions on student mentoring. This work was supported in part by the Howard Hughes Medical Institute, the National Science Foundation, and Davidson College.

REFERENCES


