Correlating Student Knowledge and Confidence Using a Graded Knowledge Survey to Assess Student Learning in a General Microbiology Classroom

Lacey Favazzo1, John D. Willford2, and Rachel M. Watson2*

1Department of Microbiology and Immunology, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642, 2Department of Molecular Biology, University of Wyoming, Laramie, WY 82071

Knowledge surveys are a type of confidence survey in which students rate their confidence in their ability to answer questions rather than answering the questions. These surveys have been discussed as a tool to evaluate student in-class or curriculum-wide learning. However, disagreement exists as to whether confidence is actually an accurate measure of knowledge. With the concomitant goals of assessing content-based learning objectives and addressing this disagreement, we present herein a pretest/posttest knowledge survey study that demonstrates a significant difference correctness on graded test questions at different levels of reported confidence in a multi-semester timeframe. Questions were organized into Bloom's taxonomy, allowing for the data collected to further provide statistical analyses on strengths and deficits in various levels of Bloom's reasoning with regard to mean correctness. Collectively, students showed increasing confidence and correctness in all levels of thought but struggled with synthesis-level questions. However, when students were only asked to rate confidence and not answer the accompanying test questions, they reported significantly higher confidence than the control group which was asked to do both. This indicates that when students do not attempt to answer questions, they have significantly greater confidence in their ability to answer those questions. Additionally, when students rate only confidence without answering the question, resolution across Bloom's levels of reasoning is lost. Based upon our findings, knowledge surveys can be an effective tool for assessment of both breadth and depth of knowledge, but may require students to answer questions in addition to rating confidence to provide the most accurate data.

INTRODUCTION

More than a century ago, Arthur Bandura began writing about the importance of self-efficacy as a predictor of behavior (4, 5). He showed that beliefs of self-efficacy could, in fact, increase or decrease performance (3). Bandura's work set the foundation for more recent educational scholars who show that student motivation, effort, persistence, and eventual academic success are impacted by confidence (12, 13, 21). Additionally, as learners' feelings of self-efficacy increase, so too does their ability to correctly self-assess (8). From this work, the confidence survey was born. Such surveys ask students to answer questions in terms of their confidence in their ability to respond rather than actually formulating an answer. Confidence surveys have been applied as educational assessment tools since the late 1990s (23). The knowledge survey is one such confidence survey that lists up to 200 content- or skill-based questions (17). Nuhfer and Knipp (17) report that, as class confidence rises, so too does class knowledge and ability. Moreover, knowledge surveys encourage highly reflective and effective pedagogical practices.

The benefits of knowledge surveys extend to students, instructors, and to the relationships between the two. Unlike traditional pre- and posttest evaluations, which address a limited amount of course material due to time and effort constraints, a much wider variety and depth of material can be covered on a knowledge survey (10, 17). Students benefit because, as there is no way to cheat on a knowledge survey, these assessments can be completed at home, in the absence of classroom pressure. Time investment is minimal as confidence ratings can be made much more rapidly than actual answers to questions can be provided. Knowledge surveys can provide a study guide to students of concepts to be covered in a class or degree program. Additionally, students can self-assess their knowledge and they are aware of the complete course or program content/expectations (6, 10). Instructors benefit because knowledge survey creation increases their preparedness and organization. They are forced to think ahead about comprehensive course content (17). Also, as students do not report confidence in concepts that have not been presented (10), knowledge surveys allow instructors to identify gaps in the curriculum.

*Corresponding author. Mailing address: Department of Molecular Biology, 1000 E. University Ave., Laramie, WY 82071. Phone: 307-766-3524. Fax: 307-766-5098. E-mail: rwatson@uwyo.edu.
†Supplemental materials available at http://jmbe.asm.org
To maximize knowledge survey effectiveness, questions should be organized into levels of cognitive thought (17), allowing students to be exposed to questions of varying difficulty and complexity (6). Cognitive learning taxonomies such as those of Anderson (1) and Bloom (7) can provide hierarchical schema to classify thought levels (19). Previous knowledge survey studies have relied on Bloom’s traditional levels of reasoning: 1) recall, 2) comprehension, 3) application, 4) analysis, 5) synthesis and 6) evaluation (10, 17). Thus, the study reported here utilized Bloom’s thought levels to categorize questions in the relative proportions as previously suggested (17).

While there is consensus in the literature that knowledge surveys have advantages, there is disagreement as to whether confidence correlates with knowledge. A study by Bowers et al. (10) showed no significant correlation between student confidence and grades. A rapid rebuttal by Nuhfer and Knipp (18) pointed out that the limits on the prior authors’ reliability did not allow for such a conclusion and, in fact, when appropriate statistics were applied to the data, a positive correlation is shown between posttest scores (graded) and knowledge surveys. Thus, as earlier argued by Nuhfer and Knipp (17), while correlations should not be made between individuals’ knowledge survey scores and grades, class confidence does seem to rise with knowledge and ability. Bowers and Brandon (9) responded to this critique with a note that many statistical methods may be appropriate and highlighted the need for additional peer-reviewed literature on the subject.

Fueled by the knowledge that surveying students’ feelings of self-efficacy could enable meaningful assessment of our learners, the interdepartmental Microbiology Program at the University of Wyoming decided, in the spring semester of 2011, to integrate knowledge surveys into our nascent assessment plan. At that time, University Assessment Coordinators had ranked the Microbiology Program as a program in need of improved evaluation of learning objectives. No efforts had been made to align course outcomes with overall program goals. The Microbiology Steering Committee set a goal of developing and implementing direct and indirect assessment of both content-based and skill- and process-based learning objectives. A curriculum map (Appendix 1) relates program learning outcomes at the onset of the project and integrates nascent assessment efforts within the coursework and learning objectives. To assess skill- and process-based objectives, a problem-based capstone course has since been designed, developed and piloted. Here we report on the development and multi-semester piloting of a knowledge survey that combined the principles of a knowledge survey with a traditional pre- and postsemester test. This study was intended to determine if knowledge surveys could have value for the Microbiology Program in direct and indirect assessment of content-based student learning objectives. Specifically, we hypothesized that there would be a correlation between student confidence and correctness with any observed correlation weakening at higher levels of Bloom’s reasoning. Furthermore, no difference was expected in the confidence level of students who had to answer the questions along with rating their confidence vs. those only asked to rate confidence. This multi-semester study also responds to Bower’s and Brandon’s (9) call for additional literature further investigating the correlation between confidence and knowledge.

**METHODS**

Development of the pretest/posttest knowledge survey and pilot semester

To develop a knowledge survey that combined the principles of a knowledge survey with a traditional pre- and postsemester test, a question database was first developed. Categories of questions were identified based on standardized post-bachelor tests, such as the GRE and MCAT, but with the primary focus on topics instructed within the Microbiology Program at the University of Wyoming. Questions were written and submitted by the authors herein and by faculty within the Microbiology Program, with students who had completed particular core microbiology courses assisting with the project. The primary goal was to obtain questions representative of those that students would encounter in the General Microbiology course in which the survey was piloted as well as other courses taken by microbiology majors. Questions were then coded into Bloom reasoning levels based upon the query sound/verb nature as previously described (17, 20). For example, questions asking students to define, list or state something (using questioning terms such as what? or who?) were categorized in the Bloom recall level. Additional Bloom levels as well as verbs/query sounds used to classify questions and sample questions can be found in Appendix 2. This information had also been distributed to faculty when requesting questions, with the hope that it would help them identify and submit questions in multiple Bloom levels.

After completing the question database, the pretest/posttest knowledge survey itself was compiled. Forty questions were chosen, with the goal of sufficiently addressing all of the classes of questions in Bloom’s reasoning without making the test too long or tedious to be taken seriously by the students. In total, 11 recall, 12 comprehension, 6 application, 5 analysis, 4 synthesis, and 2 evaluation questions were chosen to correspond roughly to the percentages in previously described knowledge surveys (17). A majority of the questions evaluated content that would specifically be covered in the General Microbiology course to which this survey was tied. Ideally, this would result in low knowledge/confidence for the pretest and higher knowledge/confidence for the posttest. Additionally, some of the questions selected were included to serve as control questions for both the pre- and posttests (i.e. material from required courses that students should know and answer correctly on both
tests or material not covered until later in their career that should be incorrect on both tests), thereby allowing for better delineation of actual knowledge and confidence in both phases of the knowledge survey. The pretest/posttest knowledge survey as administered on eCompanion can be found in Appendix 3.

Upon completion of the pretest/posttest knowledge survey, the survey, along with a statement of informed consent, was submitted to and approved by the Institutional Review Board (IRB) as Expedited Review. To pilot the pretest/posttest knowledge survey, 72 students enrolled in General Microbiology participated. This pilot was intended to establish accessibility/usability of the pretest/posttest knowledge survey itself along with the test proctoring system and format. During a week near the beginning of the semester, all of the students enrolled in the course were encouraged to take the knowledge survey. All recruitment and incentive methods were also reviewed and approved by the University’s IRB.

The pretest/posttest knowledge survey asked the students to answer the question and rate their confidence with regard to their answer. At the end of the semester, all of the students were encouraged to retake the same pretest/posttest knowledge survey and rerate their confidence on this posttest.

Pretest/posttest knowledge survey, semester 2

Having piloted the pretest/posttest knowledge survey, our goal became to collect data from additional semesters of general microbiology that recorded both students’ correctness in answering and their confidence in answering questions, to determine the correlation between class confidence and correctness of answers and to compare/contrast the correlation between class confidence and correctness of answers across Bloom’s levels of reasoning. Thus, the semester following the pilot, 110 students enrolled in general microbiology gave their informed consent to participate in the study. Any overt errors detected on the pretest/posttest knowledge survey were fixed but the only significant change from the pilot was a transition to a numerical scale (15) for confidence as opposed to the categorical scale utilized by Nuhfer and Knipp (17). This modification allowed for the confidence ranking to be more easily transformed into a continuous value with more statistical sensitivity (26). Specifically, students were instructed to mark as follows:

- “5” corresponded to the confidence that the question was answered completely for grading purposes,
- “3” indicated that students answered the question 50% correctly or knew exactly where to find the correct answer and could answer within twenty minutes, and
- “1” indicated that they were unsure of the answer or where they would find the answer.

The numbers “2” and “4” were not assigned a specific guideline for what their confidence indicated. Such graphic scales ask raters to indicate the rating by selecting an appropriate value on a horizontal line that runs from one extreme to the other. Graphic scales commonly provide specific labeled descriptions for only a subset of the valued choices, leaving the user more flexibility in interpreting unlabeled values (2). These modifications to scale and instruction (Appendix 4) were approved by the IRB as Expedited Review. In an attempt to gain the most accurate data, entry of answers and confidence into eCompanion was limited to one single confidence answer for each question. The administration of the knowledge survey was completed as described in the pilot survey.

Pretest/posttest knowledge survey, semesters 3 and 4

Following successful administration and analysis of the knowledge survey, the testing protocol was modified to evaluate the significance of requiring students to actually answer survey questions on their reported level of confidence. The knowledge survey utilized in these final two semesters (82 students and 99 students, respectively) was identical to the survey utilized in semester two. The notable modification to the administration of the knowledge survey was that after students had signed informed consent agreements, each participant was randomly assigned into one of two testing groups. Test group A was a control group and completed the pretest/posttest knowledge survey as described above (students both answered questions and rated confidence); whereas, test group B (experimental group) was only asked to register their confidence for each question of the survey. The difference in confidence for control and experimental groups was measured. For this version of the survey, final course grades (as a percentage) were also included with student information for analysis to allow for direct knowledge comparison of the groups.

Pretest/posttest grading

To promote consistency in grading, a rubric was made for each question which defined the answer required to receive full credit for the question and, if applicable, answer(s) which would receive partial credit. To simplify point distribution, every correct answer was worth 1 point and partial credit was to receive 0.5 points when awarded. For the pilot pretest, a group of student volunteers graded the questions and for all following tests, the questions were graded entirely by the lead author. For each student answer, points awarded and confidence were recorded, allowing for not only whole test analysis, but also for question-specific knowledge/confidence correlations to be drawn. Any pretest/posttest knowledge survey found during grading to be ≥ 50% incomplete was not included for further analysis.
**Data analyses**

To evaluate any correlation between student confidence and correctness, each student's response to each question, in addition to her/his confidence score for that particular question, was considered along with the average confidence and correctness on each entire test for each student. A bivariate Pearson \( r \) correlation was utilized to evaluate the correlation between confidence and correctness.

To categorize differences between the average correctness for each reported confidence level, a one-way ANOVA (\( \alpha = 0.05 \)) determining least square differences (LSD) was used to identify significant differences in correctness dependent upon confidence. In this case, confidence level was the independent variable and correctness was the dependent variable. This analysis was carried out for both the pre- and posttest each semester. Similar analyses were performed to characterize confidence and correctness according to their level of Bloom's reasoning; however, these analyses utilized Bloom's reasoning level as the independent, categorical variable and the confidence number or correctness as a dependent, numerical variable.

Confidence values of each test group from knowledge survey semesters three and four were compared as described above for both general pre- and posttest confidence as well as confidence according to Bloom's reasoning level. Mean student final course grade for each group was determined to serve as a knowledge comparison between groups.

**RESULTS**

**The knowledge survey**

The knowledge survey showed a significant difference (\( p < 0.001 \)) in mean graded correctness for each level of reported confidence. Correctness averages by confidence data from all three semesters are summarized in Table 1 and are depicted in Figure 1. It is noteworthy that while both the pre- and posttests showed differences in knowledge dependent upon confidence, the posttest differences were stronger in the higher levels of confidence as there was only a trending difference (\( p = 0.059 \)) between confidence levels 4 and 5 in the pretest data. However, a significant difference between these confidence levels was observed (\( p < 0.001 \)) in the posttest data. These data are further supported by the correlation analyses which demonstrated a moderate positive correlation in analyzing the correlation between confidence and correctness on an individual question (\( r = 0.645 \) pretest, \( r = 0.600 \) posttest) as well as a student's average confidence and average correctness across a whole exam (\( r = 0.615 \) pretest, \( r = 0.612 \) posttest). The student average data for these correlations are depicted in Figures 2 and 3. It is important to note that, as group B did not have to answer questions, none of their data were able to be included in the confidence vs. correctness analyses.

**Confidence by test group**

The format of the knowledge survey with regard to asking students to answer questions or not had a significant (\( p < 0.001 \)) impact on the confidence reported. The B group that did not have to answer questions on the pretest/posttest knowledge survey defined a confidence 27% higher on the pretest and 31% higher on the posttest than the A group. Utilizing the final grade percentage to compare the knowledge of the groups demonstrated no significant difference (\( p = 0.109 \)) in group knowledge. Moreover, while not significant, the final course average of group A was 3.5% higher than that of group B. These data are depicted in Figure 4. Despite the impact on confidence, the B test format was successful in requiring less student time on task. The A test format required significantly (\( p < 0.001 \)) more time (92.4 ± 5.7 minutes) to complete than the B test format (15.8 ± 5.1 minutes).

**TABLE 1.** Average graded correctness for each reported confidence level of the knowledge survey.

| Confidence Level | Graded Correctness | Pretest ¹ | Posttest ² |
|------------------|--------------------|----------|-----------|
| 1                | 2.7% a             | 3.0% a   |
| 2                | 20.0% b            | 21.9% b  |
| 3                | 38.1% c            | 38.2% c  |
| 4                | 61.4% d            | 61.7% d  |
| 5                | 66.3% d            | 73.1% e  |

¹Standard error = 1.4%.
²Standard error = 1.1%.
abcde Significantly different subsets exist within the pretest and posttest groups.

**FIGURE 1.** Average graded correctness for each reported confidence level. a, b, c, d, and e define significantly different subsets within the pretest (dark gray bars) and posttest (light gray bars), respectively. Reported confidence level is depicted on the X-axis with mean graded correctness on the Y-axis. Error bars define a 95% confidence interval.
Bloom’s level of reasoning

As expected, there were differences observed in mean student confidence and knowledge across different levels of Bloom’s reasoning. Students both answering the questions and rating confidence (Group A) demonstrated significantly higher \((p < 0.001)\) knowledge and confidence at the recall level (Level 1) than all other levels of Bloom’s reasoning in both the pre- and posttest, with the exception of the evaluation level (Level 6) on the posttest \((p > 0.124)\). Other than this one occurrence, the knowledge and confidence differences between Bloom’s levels of reasoning were more clearly delineated in the posttest than the pretest data. Specifically, the comprehension (Level 2), application (Level 3), and analysis (Level 4) levels showed no significant difference \((p > 0.516)\) in confidence on the pretest. However, as noted above, on the posttest, confidence in each of the first five levels of reasoning was significantly different \((p < 0.018)\) from the others. All of these data are summarized in Figure 5. Outside of comparing Bloom’s Levels 2 to 4 on the posttest, which showed similar knowledge \((p > 0.148)\), knowledge level mirrors that of confidence in each level of reasoning for the pre- and posttests. Students rating only confidence (Group B test format) did not produce similar patterns of confidence to the A test format as less of a difference in average confidences on both the pre- and posttests was measured in the B test format. Figure 6 summarizes the posttest confidence by Bloom’s level of reasoning for each group to demonstrate this difference.

DISCUSSION

The positive correlation between confidence and correctness on both pre- and posttests as well as the ability to track improvements in confidence and correctness across Bloom’s reasoning levels seen in spring 2011 pilot data fueled further exploration of the pretest/posttest knowledge...
survey. Transitioning from the “A” “B” “C” scale responses suggested by Nuhfer and Knipp (17) to a graphic numeric rating scale (2) with more confidence options enabled us to improve instrument sensitivity (26) and allowed students to more precisely relate their feelings of efficacy.

Over the three semesters that followed (fall 2011, spring 2012, and fall 2012), the consistent significant differences observed in average correctness dependent upon student confidence for both pre- and posttests indicated that students who chose a higher confidence level were significantly more likely to respond correctly. The moderate positive correlation observed between knowledge and confidence both on the individual question and entire exam level further supports this indication. These findings agree with Nuhfer and Knipp (17) in that they indicate that confidence rises with knowledge. In addition to knowledge level, increases in reported confidence may also demonstrate increases in other valuable indicators of success (motivation, effort, and persistence) (12, 13, 21) not discernible in stand-alone knowledge assessments. The less clearly defined differences observed between confidence and correctness seen on the pretests may be explained by the idea that incoming sophomores in the microbiology class do not know exactly what they do not know, particularly at higher levels of reasoning. Overconfidence in pretest knowledge surveys is well-documented and can be considered more a deficit in the ability to self-evaluate than a lack of correlation between correctness and confidence (6, 11). Additionally, the clearer observable differences in knowledge by confidence level on the posttest supports earlier work indicating that ability to accurately self-assess increases as feelings of self-efficacy increase (8).

The knowledge survey data collected provided statistical analysis on strengths and deficits in various levels of Bloom’s reasoning with regard to mean correctness. As was seen in both the pretest and the posttest in all three semesters, synthesis is a Bloom category with which learners experienced relative difficulty. The knowledge survey served to highlight this relative struggle and bring it to the attention of both students and instructors. The recall and comprehension categories, on the other hand, showed a notably significant increase in confidence and correctness from pre- to posttest, indicating that it is an area in the curriculum that is fostering feelings of self-efficacy and demonstrable student learning. The comparative decrease in improvement on higher Bloom levels could indicate a low level–thought skew in semester-long assessment measures as exams tend to focus on lower levels of cognitive thought (14). On another note, if in fact students must master one level of cognitive thought before the next (19), then the data obtained in this study may be expected. That is, this is a sophomore-level class and, on average, students may still be consumed by learning at recall and comprehension levels. This, then, would support the programmatic-wide adoption of a pretest/posttest knowledge survey as comparisons between lower and upper division courses would be expected to elucidate the gaining of higher levels of reasoning with time.

Another possible explanation for the surprisingly low scores and confidences for synthesis-level questions lies in the Bloom reasoning categories themselves. More recent educational research reverses the evaluation and synthesis categories in Bloom’s taxonomy, indicating that synthesis (creating) is at the peak, and is therefore the most difficult (1). This would help explain the results obtained regarding synthesis–versus evaluation-level questions, as consistently higher-than-expected knowledge and confidence were observed in the evaluation questions on both knowledge surveys. However, this possibility should not be overstated as it is also possible that this is due to the small number of questions for the evaluation category, which would create a statistical aberration if students had good knowledge for just one of the questions. Thus, it may be advisable to increase either the overall number of questions on the knowledge survey or to modify the Bloom level percentages, increasing higher-level questions compared with Nuhfer and Knipp’s suggestion (17).

Results of the experimental research study added during the fall of 2011 and the spring of 2012 failed to support the hypothesis that there would be no statistical difference between confidence expressed by students both answering questions and rating confidence (Group A) vs. those only rating confidence (Group B). In fact, the significant increase in confidence of the group only rating confidence indicates that when students do not attempt to answer questions, they have significantly greater confidence in their ability to answer those questions. Thus, when measured alone, confidence is inflated. This finding may help to reconcile the disagreement between Nuhfer and Knipp (17) and Bowers et al. (10) in that confidence alone, without the need for actual answering, may fail to be a good measure of knowledge. It is
also noteworthy that when students rate confidence only, resolution across Bloom's levels of reasoning is lost as is clearly visible in Figure 6. This seems to indicate that when students are not faced with actually answering the question, they are less likely to recognize the level of critical thought that will be needed. The phenomenon of overconfidence is well documented and continues to be a source of debate not only in education but in fields ranging from law to finance (16, 22, 24, 27). Informative to our findings, however, are several studies that indicate that a requirement to make explicit choices or to provide reasons for choices decreases overconfidence (15, 25). Our study seems to further support these findings.

Overall, data obtained from this initial knowledge survey study showed the promise of knowledge surveys in monitoring microbiology program content-based learning objectives. Our findings further indicate that if over-estimated confidence is an undesirable limitation, then students could be asked to both answer the questions and rate their confidence. Due to the significant differences in correctness at different levels of confidence repeatedly shown with this methodology, instructors could forgo grading the questions and rely on confidence as an accurate reflection of knowledge. However, this negates the student benefit noted by Bell (6) of decreased time to complete the assessment. Our data echo this negative aspect as students both answering the question and rating confidence spent an average of six times longer on the assessment. Further studies could examine the impact on confidence of formally informing students that they will have to answer the actual questions in their current or future coursework. Also studies could be proposed that monitor the impact of asking students to actually answer just a few of the questions on the knowledge survey. Perhaps this would lower overconfidence for all survey questions.

This project leaves us optimistic about the possible establishment of a knowledge survey platform to evaluate programmatic objectives. Upon being informed of the results of the current study, faculty program-wide have been convinced to submit questions representative of all of the individual required major courses. This has enabled the construction of a new pretest/posttest knowledge survey that more fully encompasses required-course content in pursuit of a microbiology degree. We are currently implementing this programmatic pretest/posttest knowledge survey across more courses in the microbiology program (Appendix 1). In addition to enabling more complete assessment of content-based learning objectives, the very act of question submission from all faculty encourages reflective teaching practice and is allowing our program to begin to better align content-based learning objectives with individual course objectives and to gain a common curriculum goal.

**Supplemental Materials**

Appendix 1: University of Wyoming Microbiology Program Curriculum Map

Appendix 2: Bloom's levels used to classify pretest/posttest knowledge survey questions

Appendix 3: Pretest/posttest knowledge survey

Appendix 4: Pretest/posttest knowledge survey instructions for fall 2011

**Acknowledgments**

We would like to acknowledge the Ellbogen Center for Teaching and Learning for their support through the Assessment Academy Grant. Additional thanks to Erika Prager, Assessment Specialist, and the Microbiology Steering Committee (Program Head, Gerard Andrews) for support in this project. Student assistants for this project were pivotal: Leslie Graul, Becca Shahi, Beau Stricker, and Chelsea Braun. The authors declare that there are no conflicts of interest.

**References**

1. Anderson, L., and D. Krathwohl. 2000. A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives. Pearson, London.

2. Ary, D., L. C. Jacobs, A. Razavieh, and C. Sorensen. 2006. Introduction to research in education, 7th ed, Vol. 1. Thomson-Wadsworth, Belmont, CA.

3. Bandura, A. 1989. Regulation of cognitive processes through perceived self-efficacy. Dev. Psychol. 25:729–735.

4. Bandura, A. 1977. Self-efficacy: towards a unifying theory of behavioral change. Psychol. Rev. 84:191–215.

5. Bandura, A. 1986. Social foundation of thought and action. Englewood Cliffs, London, UK.

6. Bell, P., and D. Volckmann. 2001. Knowledge surveys in general chemistry: confidence, overconfidence, and performance. J. Chem. Educ. 88:1469–1476.

7. Bloom, B. S., M. D. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl. 1956. Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. David McKay Company, New York, NY.

8. Bouffard-Bouchard, T., S. Parent, and S. Larivée. 1999. Influence of self-efficacy on self-regulation and performance among junior and senior high-school aged students. Int. J. Behav. Develop. 14:153–164.

9. Bowers, N., and M. Brandon. 2006. Response: re: the use of a knowledge survey as an indicator of student learning in an introductory biology course. Cell Biol. Educ. 5:315.

10. Bowers, N., M. Brandon, and C. D. Hill. 2005. The use of a knowledge survey as an indicator of student learning in an introductory biology course. Cell Biol. Educ. 4:311–322.

11. Clauss, J., and K. Geedey. 2010. Knowledge surveys: students’ ability to self-assess. J. Sch. Teach. Learn. 10:14–24.

12. Collins, J. L. 1982. Self-efficacy and ability in achievement behavior. Annual Meeting of the American Educational Research Association. New York, NY.

13. Crozier, R. 1997. Individual learners: personality differences in education. Routledge, London.
14. Gardiner, L. 1994. Redesigning Higher Education: Producing Dramatic Gains in Student Learning. Higher Education Report No. 7. ASHE-ERIC, Washington, DC.
15. Koriat, A., S. Lichtenstein, and B. Fischhoff. 1980. Reasons for confidence. J. Exper. Psych. Human Learn. Mem. 6:107–118.
16. Lichtenstein, S., B. Fischhoff, and L. D. Philips. 1982. Calibration of probabilities: the state of the art in 1980. In Kahneman, D., Slavic, P., and Tversky, A. (ed.), Judgement Under Uncertainty: Heuristics and Biases. Cambridge University Press, Cambridge, UK.
17. Nuhfer, E., and D. Knipp. 2003. The knowledge survey: a tool for all reasons. To Improve the Academy 21:59–78.
18. Nuhfer, E., and D. Knipp. 2006. Re: The use of a knowledge survey as an indicator of student learning in an introductory biology course. Cell Biol. Educ. 5:313–314.
19. O’Neill, G., and F. Murphy. 2010. Guide to taxonomies of learning. UCD Teaching and Learning Assessment Resources. [Online.] Accessed May 2014 at http://www.ucd.ie/t4cms/ucdtla0034.pdf.
20. Overbaugh, R. C., and L. Schultz. n.d. Bloom’s Taxonomy. [Online.] Accessed August 2009 at http://ww2.odu.edu/educ/roverbau/Bloom/blooms_taxonomy.htm.
21. Pajares, F. 2002. Gender and perceived self-efficacy in self-regulated learning. Theory Pract. 41:116–125.
22. Ronis, D. L., and J. F. Yates. 1987. Components of probability judgement accuracy: individual consistency and effects of subject matter and assessment method. Organiz. Behav. Human Decision Proc. 40:193–218.
23. Sander, P., and L. Sanders. 2007. Measuring confidence in academic study: a summary report. J. Res. Educ. Psych. Ped. 1:1–17.
24. Sieck, W. R., E. C. Merkle, and T. Van Zandt. 2007. Option fixation: a cognitive contributor to overconfidence. Organiz. Behav. Human Decision Proc. 103:68–83.
25. Sniezek, J. A., P. W. Paese, and F. S. Switzer III. 1990. The effect of choosing on confidence in choice. Organiz. Behav. Human Decision Proc. 46:264–282.
26. Warner, R. M. 2008. Applied Statistics: from bivariate through multivariate techniques. Sage Publications, Inc., Thousand Oaks, CA.
27. Wright, G. 1982. Changes in realism of probability assessments as a function of question type. Acta Psychologica 52:165–174.