# *Article*

# Evaluation of the Redesign of an Undergraduate Cell Biology Course

**Laura April McEwen,\* dik Harris,[†] Richard F. Schmid,[‡] Jackie Vogel,[§]
Tamara Western,[§] and Paul Harrison[§]**

*Tomlinson University Science Teaching Project, McGill University, and Faculty of Education, Queens
University, Kingston, Ontario, Canada K7M 5R7; [†]Tomlinson University Science Teaching Project, McGill
University, and Department of Physics, McGill University, Montreal, Quebec, Canada H3A 2T8; [§]Department
of Biology, McGill University, Montreal, Quebec, Canada H3A 1B1; and [‡]Department of Education, Concordia
University, Montreal, Quebec, Canada H3G 1M8

This article offers a case study of the evaluation of a redesigned and redeveloped laboratory-based cell biology course. The course was a compulsory element of the biology program, but the laboratory had become outdated and was inadequately equipped. With the support of a faculty-based teaching improvement project, the teaching team redesigned the course and re-equipped the laboratory, using a more learner-centered, constructivist approach. The focus of the article is on the project-supported evaluation of the redesign rather than the redesign per se. The evaluation involved aspects well beyond standard course assessments, including the gathering of self-reported data from the students concerning both the laboratory component and the technical skills associated with the course. The comparison of pre- and postdata gave valuable information to the teaching team on course design issues and skill acquisition. It is argued that the evaluation process was an effective use of the scarce resources of the teaching improvement project.

## INTRODUCTION

Improving the quality of the undergraduate learning experience is a major concern of today's universities. Because the most important elements of this experience come from the academic programs, it is natural that significant effort would be devoted to improving the teaching and learning of individual courses embedded within a cohesive curriculum. Unfortunately, the process of improving courses is usually left solely to the instructors, often working in isolation. It is thus unlikely that curriculum development is guided by the principles of instructional design, namely, the systematic analysis of the learning context, the students, and the learning outcomes, followed by a structured strategy for the design, production, delivery, and evaluation of instructional materials (Smith and Ragan, 2005, p. 10). Even when there is an intention to apply the principles, it is not unusual that local constraints dictate the omission of parts of the process. In particular, with course redesign initiatives that originate

with the instructors, it is typical that the window of opportunity for the redesign is limited to the evaluative phase (Smith and Ragan, 2005, p. 360).

Therefore, it is difficult to find science courses designed or redesigned in ways that explicitly respect these principles (Sundberg *et al.*, 1992; Handelsman *et al.*, 2004). Even when there is support from a teaching and learning center, resources for course development are usually in short supply, so that an initial appraisal may serve in lieu of help with systematic analysis, and a set of guidelines in lieu of help with the design of materials (McAlpine and Gandell, 2003). The appraisal and the guidelines typically conform to the notion of "constructive alignment" Biggs (1996, 1999), namely, a "marriage between a constructivist understanding of learning and an aligned design for teaching" (Biggs, 1999, p. 26). The most concrete representation of constructivism is embodied within the 14 American Psychological Association (APA, 1997) learner-centered principles, which emphasize that active, meaningful learning best occurs when attention is paid to the social context and the environment of the learners. Correspondingly, Biggs' notion of alignment implicates several elements: notably, curriculum objectives de-

rived from real-world problems, authentic assessment procedures, and teaching methods that support interaction between teacher and student (Biggs, 1999, p. 25). Only when all such elements are mutually coherent and supportive will there be "desired outcomes learned in a reasonably effective manner" (Biggs, 1999).

Even when the above-mentioned principles are implicitly or explicitly attended to, the question remains, Was the redesign effective? Unfortunately, it is this evaluative phase of the instructional design process that is the most neglected, despite its recognized importance (Smith and Ragan, 2005). A case in point is the concerted effort to introduce new forms of pedagogy by Entwistle (2005) and colleagues. They have undertaken an in-depth study of discipline-specific teaching and learning with a view to increasing instructors' capacity to make decisions based on data, carefully obtained and analyzed, that is, "the capacity for evidence-based practice." Their study has identified different "ways of thinking and practicing" in several different disciplines, one of which is biology (Hounsell and McCune, 2002; McCune and Hounsell, 2005), and they demonstrate how the teaching and learning of a discipline involves much more than the mere transmission and reception of factual information. Ironically, there are no published accounts of modifications to courses that result from the interventions of the Entwistle research; consequently, there is as yet no account of an evaluative phase.

Ordinarily, the evaluation of a course consists of the summative assessment of student learning, yielding grades, and the evaluation of the instructor(s), satisfying institutional requirements. Although both of these measures can be informative, they seldom address all aspects of a (re)design project. Indeed, after implementing a (re)designed course, it is important to know whether the students have achieved the goals set for them, and whether their feelings about the course match their expectations (and those of their instructors; Smith and Ragan, 2005, p. 345). In a university context, these pertinent questions are almost never asked (McAlpine and Gandell, 2003). Equally, data on the role of the course in the larger curriculum—"ways of thinking and practicing" (Entwistle, 2005)—are seldom produced (Tovar *et al.*, 1997). To address these shortcomings, this article offers a case study of an evaluation of the extent to which key components of a redesign were effectively implemented. It describes both the instruments used to capture additional types of information, and how the evaluation was carried out, by using pre- and postdata from two cohorts of students from successive semesters. An indirect outcome of the study is an appreciation of the role that external project support can play in improving the teaching/learning process in an undergraduate course.

## LABORATORY INSTRUCTION

Laboratory courses are an integral part of the undergraduate curriculum in any science program. Science instructors, whether they are in high school, college, or university, and whether they be in physics or cell and molecular biology, all agree that the laboratory is an essential component of the learning experience (Hodson, 1988; Sundberg *et al.*, 1992; Arons, 1993, p. 278). However, there is less consensus about

the objectives of the laboratory—whether they be those of the instructor or those of the student (Hodson, 1993). This lack of consensus is documented for American high schools in a National Academy of Sciences (2005) report and in several research-based surveys (Tobin, 1990; Hofstein and Lunetta, 2004; Hodson, 2005) and for European universities in the Labwork in Science Education Project (Seré, 2002).

Studies have shown that lab instructors' objectives can be grouped into three main categories (Seré, 2002): conceptual, epistemological, and procedural. Within the conceptual category, one might include the understanding of scientific concepts as well as the aspects of problem solving and critical thinking that relate to science (Hofstein and Lunetta, 2004). The epistemological category, in turn, might include the acquisition of scientific habits of mind and the appreciation of the nature of science itself (Hofstein and Lunetta, 2004). Finally, and of particular relevance to this article, the procedural category includes the particular skills of doing science in a laboratory: measurement, data processing, and experimental design (Seré, 2002).

Addressing the last category of Seré (2002), there is evidence that there is more emphasis on procedural skills in universities as opposed to high schools (Welzel *et al.*, 1998; also see Working Paper 6 of the European Project on Labwork in Science Education, www.idn.uni-bremen.de/pubs/Niedderer/1998-LSE-WP6.pdf). Equally, the same authors indicate that in the biological sciences this emphasis seems more pronounced than in the physical sciences, in which, typically, the instructors are focusing more on conceptual understanding. Presumably, in the biological sciences, the mastery of laboratory techniques is more essential for meaningful data collection and analysis. The course whose evaluation is described in this article is consistent with this generalization, because one of the teaching team's concerns was that the course becomes an effective means for the learning of essential laboratory techniques in cell biology. They were equally concerned that the students understood both the techniques and the results obtained through their use.

## THE CONTEXT

An initiative of the Tomlinson University Science Teaching Project was to fund faculty members in the Faculty of Science for teaching improvement. A condition for this funding was the inclusion of an evaluation component, which was intended to 1) document the impact on the quality of student learning and 2) provide evidence of the effective use of Tomlinson Project resources. The evaluation was carried out by members of the Tomlinson Project, who were external to the teaching team, but who worked in direct and ongoing collaboration with them.

The project began in spring 2004, when one of us (Vogel) applied to the Tomlinson Project for funding to redesign a core laboratory course offered by the Department of Biology. The course is the required cell biology laboratory for all biology students as well as for students from several biomedical departments, offered in both fall and winter semesters, with an average annual enrollment of approximately 240 students. It is composed of 12 weekly lectures (1 h) and 12 weekly laboratory sessions (6 h). It is team-taught by

three of us (Vogel, Western, and Harrison), each assuming responsibility for the portion of the course related to his or her individual expertise. Broadly, it has 5 wk of cell biology/biochemistry, 5 wk of molecular biology, and 2 wk of bioinformatics, and its primary objective is to provide students with "hands-on" experience in laboratory methods. Over time, the course had become increasingly outdated and was becoming completely disconnected from the methods commonly referenced during lectures (which were current). Worse, the prerequisite courses gave exposure to modern methods, so that the experience of the cell biology laboratory was profoundly demotivating.

The primary intention of the redesign was thus to better align the curriculum both with regard to content and to pedagogical strategies. Content issues were addressed by introducing a more logical ordering for the material, by integrating modern molecular and cellular biology methods in the laboratory exercises and by coupling these with bioinformatics methods currently in use in modern biotechnology research and laboratories. Simultaneously, attention was paid to coherence with the prerequisite courses.

To address a constructivist pedagogy, the redesign moved away from "cookbook" lab procedures. In the weekly lab session, each student was presented with his or her own scenario and was required to use a designated technique or techniques to solve the problem embedded within it. This more open-ended approach has been shown to result in greater student understanding and retention of molecular and cellular biology concepts, an enhanced appreciation of the laboratory environment, and increased motivation (Hodson, 1988; Bransford *et al.*, 2000). It was also anticipated that students would respond positively to the material environment—new equipment, including laptop computers—and to the relevance of the laboratory manuals. Such expectations are also consistent with the conclusions of a study by Breen and Lindsay (2002), where "deriving enjoyment from learning experiences is shown to be an important component of student motivation in several disciplines, including biology."

## THE SUPPORT

By the time the teaching team (Vogel, Western, and Harrison) received their funding from the Tomlinson Project (represented by McEwen and Harris), the first phase of their redesign was well advanced. They had already carried out a systematic analysis of the learning context, the learners, and the learning outcomes, drawing, in particular, upon their knowledge of the current status of their discipline with respect to their department's program. Equally, they had already established a plan for the production of instructional materials, namely, laboratory manuals and associated worksheets. Once the members of the Tomlinson Project joined the team, to assist with the evaluative phase, it was agreed that data collection should not be summative in nature but should focus on evidence for the effectiveness of the redesign as perceived by the end users—the students. It would necessarily be informed by all the preceding design steps carried out by the teaching team.

It was decided to measure effectiveness in terms of two dimensions, namely, the improvement of the students' technical skills (the primary objective of labs) and the psychosocial factors affecting the learning environment and the students' motivation. The first of these dimensions addresses the procedural category of Seré (2002), and the second dimension addresses certain aspects of her epistemological and conceptual categories.

For the psychosocial dimension, the chosen instrument was the Science Laboratory Environment Inventory (SLEI) (Fraser *et al.*, 1995), which was validated in university classes in Australia and other countries. The authors of this instrument suggest that it can be used to "monitor students' views of their laboratory classes, to investigate the impact that different laboratory environments have on student outcomes, and to provide a basis for guiding systematic attempts to improve these learning environments" (Fraser *et al.*, 1995, p. 415). The SLEI measures the psychosocial aspects of laboratory learning contexts on five subscales: cohesiveness (C), open-endedness (O), integration (I), rule clarity (R), and material environment (M). Definitions of the subscales are presented in Table 1. Each subscale has seven questions, each answered on a Likert scale from 1 to 5, with 5 representing the most positive response. This instrument addresses several aspects associated with the APA principles. Additionally, the integration subscale directly addresses constructive alignment in the sense of Biggs (1996, 1999).

For the technical skills, the curriculum skills matrix of Caldwell *et al.* (2004) was used: It was specifically designed for developing and assessing undergraduate biochemistry and molecular biology laboratory curricula. Students were asked to document their level of familiarity with the course's 20 laboratory techniques identified by the teaching team as core to the curriculum skills matrix. The first column of Table 2 lists these techniques, whereas the second column indicates where in the course they were introduced. The third column identifies certain skills that had been encountered previously by the students, and with which, according to laboratory staff, they might be expected to be more or less familiar.

**Table 1.** Definitions of SLEI subscales

| Subscale[a] | Definition |
|---|---|
| C | Extent to which students know, help, and are friendly toward each other |
| O | Extent to which the laboratory activities emphasize an open-ended divergent approach to experiment |
| I | Extent to which the laboratory activities are integrated with nonlaboratory and theory classes |
| R | Extent to which behavior in the laboratory is guided by formal rules |
| M | Extent to which the laboratory equipment and materials are adequate |

[a] C, cohesiveness; O, open-endedness; I, integration; R, rule clarity, and M, material environment

Adapted from Fraser *et al.* (1995)

## DATA COLLECTION AND ANALYSIS

In principle, important information could have been obtained from comparisons with so-called control groups. However, in the present situation, it would have been ethically undesirable to maintain two cohorts of students registered in similar degree programs at the same point in time with one group following an openly recognized outdated curriculum (a practice that the department would likely not have accepted). And, for past cohorts, there were no data with which a comparison could be made.

Thus, a pretest/posttest (repeated measures) design was implemented for the evaluation. In the first week of classes, students were asked to document their expectations about aspects of the laboratory environment and their familiarity with the 20 targeted technical skills. Then, they reported on their actual laboratory experiences and their enhanced familiarity with the technical skills at the end of the semester. The same protocol was implemented for both fall 2004 and winter 2005 cohorts. Consent forms and complete data were collected on 76 students in the fall 2004 semester and on 93

students in winter 2005, representing approximately 60 and 70% participation rates, respectively.

### SLEI Results

For the psychosocial, laboratory environment aspects, individual subscale scores were calculated for each student by summing their responses for the seven questions in each of the five subscales of the SLEI, giving a maximum possible score of 35. Pre- and posttest global mean scores were then calculated for the two cohorts: These scores are presented in Figure 1 as four bars for each subscale.

The results, for all except the open-endedness subscale, reveal that students' expectations were initially very high: global mean scores on the pretest ranged from 28.7 to 33.3 in fall 2004 and from 29.7 to 33.2 in winter 2005. The posttest scores ranged from 25.8 to 30.3 in fall 2004 and from 25.5 to 28.5 in winter 2005. Parametric statistical analyses ($t$ tests) revealed significant decreases, flagging a potential source of concern—that the redesign had produced only qualified success.

However, there are two unrelated reasons permitting a more positive interpretation. First, there is previous, related research that has found similar statistically significant decreases in situations where the classroom environment has been measured both before and after improvements have been implemented (Fraser, 1998), suggesting that such decreases are to be expected and should not necessarily be causes for concern. For example, in the particular study of a science classroom reported by Fraser and Fisher (1986), using not the SLEI but the more generic Classroom Environment Survey, the students' actual environment scored significantly ($p < 0.05$) lower than their preferred environment, namely, their expectations, on three of the six subscales. Second, there is a statistical argument based on the observation that the distributions of scores for the different subscales are highly and negatively skewed, corresponding to mean scores that are very high. In such a situation, the usual interpretation of parametric analyses, designed for normal distributions, is questionable (Coladarci *et al.*, 2008).

The students' uncharacteristically high expectations may be attributed to their anticipation of the features of a new course, with a completely new, computerized lab. In this context, mean posttest scores that dropped less than half a point on the 5-point scale still place outcomes as highly

**Table 2.** Technical skills matrix

| Technique | Exposure | Status at entry |
|---|---|---|
| 1. Centrifugation | All wet labs | Previously encountered |
| 2. Pipetting (traditional and micropipettor) | All wet labs | Previously encountered |
| 3. Purification of DNA (genomic and plasmid) | Labs 1 and 4 | |
| 4. Agarose gel electrophoresis | Labs 2–5 | Previously encountered but prepared for them |
| 5. Polymerase chain reaction (PCR) | Labs 2 and 5 | |
| 6. DNA cloning (restriction digests, ligation, and transformation) | Labs 3 and 4 | |
| 7. Growing bacteria | Lab 3 and 4 | Previously encountered/could have done it |
| 8. Reverse transcription-PCR | Lab 5 | |
| 9. Analysis of promoter–GUS fusions | Lab 5 | |
| 10. Light microscopy | Lab 6 | Previously encountered and used |
| 11. Protein purification | Lab 7 | |
| 12. SDS-page | Labs 11–14 | |
| 13. 2D-page | Lab 9 | |
| 14. Immunoblot | Lab 10 | |
| 15. Protein and gene sequence comparisons | Labs 11 and 12 | |
| 16. BLAST | Labs 11 and 12 | |
| 17. ExPASy tool use | Labs 11 and 12 | |
| 18. Motif searches | Labs 11 and 12 | |
| 19. Structure prediction | Labs 11 and 12 | |
| 20. Function prediction | Labs 11 and 12 | |

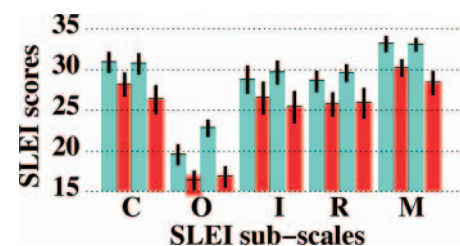Selected and adapted from Caldwell *et al.* (2004)



**Figure 1.** Pre- and post-SLEI subscale scores, indexed as in Table 1, with four bars for each subscale. From left to right, the bars represent the pre- and posttests for the fall 2004 cohort and pre- and posttests for the winter 2005 cohort, respectively. Pretest results are shown in blue, and posttest results are shown in red. Centered on the top of each bar are vertical lines, whose length represents 1 SD.

positive. The integration subscale scores were particularly gratifying, with posttest means hovering at 4 of 5 in an area traditionally problematic in laboratory courses, namely, the disconnect between lectures and laboratories. That the material environment subscale dropped the most, both in fall 2004 and winter 2005, is also consistent with this interpretation, because pedagogically sound technology applications seldom match the expectations of "digital natives" (Prensky, 2001).

Results from the open-endedness subscale differed from the rest. Although the other four subscales yield highly skewed, positive attitudes initially, students had much lower expectations for open-endedness in both semesters, with pretest scores of 19.6 and 22.9 in fall 2004 and winter 2005, respectively, and even lower scores, 16.5 and 16.9, respectively, on the posttests. These low scores are entirely consistent with the original data of Fraser *et al.* (1995), obtained from posttests administered to both high school and university students. This study discussed "an international pattern in which science laboratory classes . . . are dominated by closed-ended activities." Such prior experiences of the cell biology students presumably explain their pretest scores.

However, even in this context, the even lower posttest scores reported by both cohorts were disappointing in view of the attempt to provide more open-ended activities via the scenario format. A possible explanation comes from examining the results of a study by Martin-Dunlop and Fraser (2007). In this case, posttest scores on the open-endedness subscale exceeded students' expectations. The emphasis of this course was not on coverage of material but rather on the provision of support and personalized guidance for students in their open-ended activities, achieved by dividing a class of 250 students into laboratory sections of approximately 25, each with an experienced instructor. Although future cell biology courses may approach such a format by appropriate preparation and deployment of the teaching assistants, the suspicion is that the open-endedness will always be compromised by the competing requirement to "cover" all the necessary skills—a common dilemma in laboratory classes.

### Technical Skills Results

Mean scores were calculated for each of the 20 targeted technical skills for both cohorts. They are displayed in Figure 2, as groups of four bars, each group corresponding to a single technical skill, indexed as in Table 2.

Before elaborating on these results, it should be noted that in fall 2004, the students were provided four response options: 1) "heard of technique," 2) "understand technique," 3) "have performed technique," and 4) "comfortable performing technique," and these options were allocated scores of 1, 2, 3, and 4, respectively. However, this set of options did not anticipate the possibility that some students would claim to have no knowledge of certain targeted technical skills. Consequently, the cohort in winter 2005 was offered a fifth alternative, "never heard of technique," which was allocated a score of zero. The pretest scores for fall 2004 and winter 2005 are therefore not strictly comparable, although the posttest scores, when scores of zero are impossible, are comparable—and, indeed, remarkably similar. The statistical analyses below, however, compare only pretest with

posttest for fall 2004 and winter 2005 separately: they are not compromised by the differing scales.

Results show that students in both cohorts began the course with similar prior knowledge and experience of specifically targeted technical skills. Although the majority of students reported being "comfortable performing techniques" 1 and 2 (centrifugation and pipetting), most reported only "an understanding" of techniques 4, 7, and 10 (electrophoresis, growing bacteria, and microscopy), and most had only "heard of" the remaining 15 techniques. Indeed, the majority of the fall 2005 cohort reported that techniques 9, 13, and 17 (promoter–$\beta$-glucuronidase [GUS] fusions, two-dimensional polyacrylamide gel electrophoresis [2D-PAGE], and Expert Protein Analysis System [ExPASy]) were "never heard of" and that many other techniques were either "never heard of" or merely "heard of." After the course, however, both cohorts reported a significant degree of comfort with all techniques, including 9, 13, and 17, with techniques 1 and 2 approaching very comfortable. The data thus indicate that the course produced major improvements in students' ability to understand and perform current laboratory techniques. Indeed, all 20 correlated $t$ tests for both cohorts yielded highly significant increases, even for the techniques associated with significant prior knowledge (techniques 1, 2, 4, 7 and 10). In all cases, the $t$ tests yielded $p < 0.01$, save for technique 10, winter term, which gave $p < 0.015$.

## DISCUSSION AND CONCLUSIONS

The primary focus of this article has been on the evaluation of a redesigned course made possible via resources from a teaching improvement project. Thus, the discussion ad-
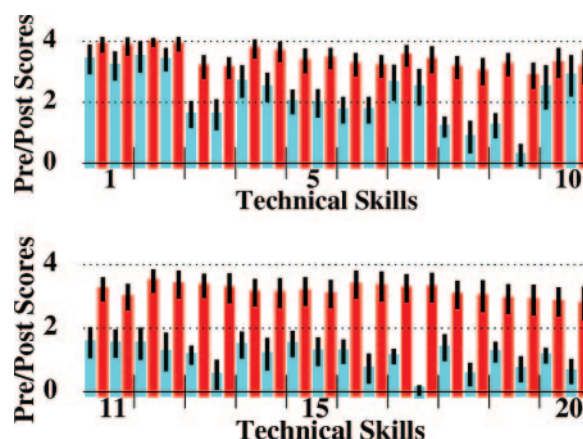


**Figure 2.** Pre- and postscores for technical skills 1–9 (molecular biology) and 10 (microscopy), indexed as in Table 2, with four bars corresponding to each skill. Within each group, from left to right, the bars represent the pre- and posttests for the fall 2004 cohort and the pre- and posttests for the winter 2005 cohort, respectively. Pretest results are shown in blue, and posttest results are shown in red. Centered on the top of each bar is a vertical line whose length represents 1 SD. Pre- and postscores for technical skills 11–14 (protein analysis) and 15–20 (bioinformatics), indexed as in Table 2, with four bars corresponding to each skill. Detailed features of this figure are as described for technical skills 1–10.

dresses how the project's support for the evaluative phase provided evidence that the redesign achieved its stated learning goals and provided benefit to the teaching team.

The support for the evaluation described in this article influenced the implementation of the redesign in two ways. First, as noted above, and examined by research on reflective practice (Schön 1995), teaching is often a very private process, with instructors working in isolation. The evaluation prompted a collaborative process between a teaching team and a group of colleagues from outside the discipline. What emerged was the need for the articulation of a process that is usually left unstated. External criteria had to be established, and the means to assess them determined as a joint, cooperative exercise. The second outcome, connected with the first outcome, was that the learner-centered approach necessitated the use of validated instruments to give the teaching team direct feedback on two aspects of their redesign: improvement of the laboratory environment and greater familiarity with certain technical skills. Although traditional summative assessment of student achievement at the end of the course is essential, these instruments provided additional, more direct information about the validity and impact of curricular decisions. Although comparisons with data from control groups or from past cohorts was not possible, future research, using the same instruments, could examine constructive alignment within the revised course design to better understand how evaluation impacts reflective practice.

In large measure, the principles of instructional design were observed during the course redesign process. This was the case even in the initial planning phase, before the involvement of the Tomlinson Project. For example, a developmental sequence was built into the curriculum design (proceeding from simple to complex skills), and implementation seemed to be effective. Explicitly, the design principles determined the nature of the evaluative phase and the choice of the SLEI instrument as a context-specific measure of constructive alignment. Although the results suggested that the students did not fully appreciate the redesigned learning environment, their posttest data still represented a very high level of satisfaction. Furthermore, they did report significant gains in the mastery of all the measured techniques. If there was a shortcoming, it was that the evaluation did not adequately examine the nature and extent of the transfer of conceptual content to the laboratory from the accompanying lectures. Although the integration index provided an indirect measure of this phenomenon, it would have been valuable to probe more deeply the extent of this form of alignment.

Finally, what was the benefit to the teaching team? First was evidence that the redesign led to a high degree of student satisfaction. But more importantly, perhaps, was an appreciation that support from colleagues outside the discipline could be pertinent and valuable. The identification and adaptation of the SLEI and the curriculum skills matrix were central to the evaluative phase, demonstrating that contributions from the educational literature can play a crucial role. Interventions in course redesign from those who have and claim no discipline-specific expertise are often viewed with suspicion, if not with outright rejection. This study serves as an example of where value is added due to "outsider's perspectives."

From the viewpoint of the Tomlinson Project, was the decision to provide support only in the form of an evaluation of the redesign an effective use of resources? After all, the evaluation only began after the instructional design process was already well advanced. This late entry point meant that it was not possible to compare either student satisfaction or student outcomes with those for the previous version of the course. Intervention at an earlier stage of the redesign might conceivably have resulted in the identification of different, possibly more explicit, learning outcomes. Would these have been more pertinent in terms of the students' learning experience? It is impossible to say. However, involvement at a later stage is not unusual in projects of this type (Smith and Ragan, 2005) and, in this instance, it provided evidence for the positive outcomes of the redesign that would otherwise not have been available. Such evidence was important because it documented not only the improvement in student learning but also the effective deployment of resources from the Tomlinson Project. In the context of limited resources, the choice to focus on evaluation is therefore a clear advantage, because other choices can lead to the absence of evidence for effectiveness (McAlpine and Gandell, 2003).

## ACKNOWLEDGMENTS

## REFERENCES

Arons, A. B. (1993). Guiding insight and inquiry in the introductory physics laboratory. Phys. Teach. *31*, 278–282.

American Psychological Association (1997). Learner-centered Psychological Principles: A Framework for School Redesign and Reform. http://www.apa.org/ed/lcp2/lcp14.html (accessed 9 January 2009).

Biggs, J. B. (1996). Enhancing teaching through constructive alignment. High. Educ. *32*, 347–364.

Biggs, J. B. (1999). Teaching for Quality Learning at University, Buckingham, United Kingdom: Society for Research into Higher Education and Open University Press.

Bransford, J. D., Brown, A. L., and Cocking, R. R. (2000). How People Learn: Brain, Mind, Experience, and School, Washington, DC: National Academies Press.

Breen, R., and Lindsay, R. (2002). Different disciplines require different motivations for student success. Res. High. Educ. *43*, 693–725.

Caldwell, B., Rohlman, C., and Benore-Parsons, M. (2004). A curriculum skills matrix for development and assessment of undergraduate biochemistry and molecular biology laboratory programs. Biochem. Mol. Biol. Educ. *32*, 11–16.

Coladarci, T., Cobb, C. D., Minium, E. W., and Clarke, R. C. (2008). Fundamentals of Statistical Reasoning in Education, New York: Wiley.

Entwistle, N. (2005). Ways of Thinking and Ways of Teaching across Contrasting Subject Areas. Paper prepared for the ISL 2005 Conference, Improving Student Learning by Assessment, 5–7 September 2005, London, United Kingdom. http://www.ed.ac.uk/etl/publications.html (accessed 9 January 2009).

Fraser, B. J., and Fisher, D. L. (1986). Using short forms of classroom climate instruments to assess and improve classroom psychosocial environment. J. Res. Sci. Teach. *23*, 387–413.

Fraser, B. J., Giddings, G. J., and McRobbie, C. J. (1995). Evolution and validation of a personal form of an instrument for assessing science laboratory classroom environments. J. Res. Sci. Teach. *32*, 399–422.

Fraser, B. J. (1998). Science Learning Environments: Assessment, Effects and Determinants. In: International Handbook of Science Education (ed. B. J. Fraser and K. G. Tobin), Dordrecht, The Netherlands: Kluwer, 527–564.

Handelsman, J., *et al*. (2004). Scientific teaching. Science *304*, 521–522.

Hodson, D. (1988). Experiments in science and science teaching. Educ. Philos. Theor. *20*, 53–66.

Hodson, D. (1993). Re-thinking old ways: towards a more critical approach to practical work in school science. Stud. Sci. Educ. *22*, 85–142.

Hodson, D. (2005). Teaching and learning chemistry in the laboratory: a critical look at the research. Educ. Quim. *16*, 30–38.

Hofstein, A., and Lunetta, V. N. (2004). The laboratory in science education: foundations for the Twenty-First Century. Sci. Educ. *88*, 28–54.

Hounsell, D., and McCune, V. (2002). Teaching-Learning Environments in Undergraduate Biology: Initial Perspectives and Findings. ETL Occasional Report 2. http://www.ed.ac.uk/etl/publications.html (accessed 9 January 2009).

Martin-Dunlop, C., and Fraser, B. J. (2007). Learning environment and attitudes associated with an innovative science course designed for prospective elementary teachers. Int. J. Sci. Math. Educ. *6*, 163–190.

McCune, V., and Hounsell, D. (2005). The development of students' ways of thinking and practicing in three final-year biology courses. High. Educ. *49*, 255–289.

McAlpine, L., and Gandell, T. (2003). Teaching improvement grants: their potential to promote a scholarly approach to teaching. J. Further High. Educ. *27*, 187–194.

National Academy of Sciences (2005). America's Lab Report: Investigations in High School Science. www.nap.edu/catalog.php?record_id=11311 (accessed 9 January 2009).

Prensky, M. (2001). Digital natives; digital immigrants. On the Horizon *9*, 1–9. www.scribd.com/doc/9800/Prensky-Digital-Natives-Digital-Immigrants-Part 2.

Seré, M. G. (2002). Towards renewed research questions from the outcomes of the European project labwork in science education. Sci. Educ. *86*, 624–644.

Schön, D. (1995). The new scholarship requires a new epistemology. Change *27*, 27–34.

Smith, P. L., and Ragan, T. J. (2005). Instructional Design, San Francisco, CA: Wiley/Jossey-Bass.

Sundberg, M. D., Kormondy, E. J., Carter, J. L., Moore, J. A., Postlethwait, S. N., and Thornton, J. W. (1992). Reassessing the Commission on Undergraduate Education in the Biological Sciences. Bioscience *42*, 442–447.

Tobin, K. (1990). Research on science laboratory activities: in pursuit of better questions and answers to improve learning. Sch. Sci. Math. *90*, 403–418.

Tovar, M., Gagnon, F., and Schmid, R. F. (1997). Development of a consultation profile of interventions perceived as successful by human performance technology consultants. Perform. Improv. Q. *10*, 67–83.

Welzel, M., Haller, K., Bandiera, M., Hammelev, D., Kouramas, P., Niedderer, H., Paulsen, A., Robinault, K., and Aufschnaiter, S. V. (1998). Ziele, die Lehrende mit experimentellen Arbeiten in der naturwissenschaftlichen Ausbildung verbinden. Z. Didaktik Naturwiss. *4*, 29–44.