A set of vertically integrated inquiry-based practical curricula that develop scientific thinking skills for large cohorts of undergraduate students

Kirsten Zimbardi,1 Andrea Bugarcic,1,2 Kay Colthorpe,1 Jonathan P. Good,3 and Lesley J. Lluka1

1School of Biomedical Science, The University of Queensland, Brisbane, Queensland, Australia; 2Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia; and 3Department of Natural Sciences, Daemen College, Amherst, New York

Submitted 6 June 2012; accepted in final form 28 September 2013

Zimbardi K, Bugarcic A, Colthorpe K, Good JP, Lluka LJ. A set of vertically integrated inquiry-based practical curricula that develop scientific thinking skills for large cohorts of undergraduate students. Adv Physiol Educ 37: 303–315, 2013; doi:10.1152/advan.00082.2012.—Science graduates require critical thinking skills to deal with the complex problems they will face in their 21st century workplaces. Inquiry-based curricula can provide students with the opportunities to develop such critical thinking skills; however, evidence suggests that an inappropriate level of autonomy provided to underprepared students may not only be daunting to students but also detrimental to their learning. After a major review of the Bachelor of Science, we developed, implemented, and evaluated a series of three vertically integrated courses with inquiry-style laboratory practicals for early-stage undergraduate students in biomedical science. These practical curricula were designed so that students would work with increasing autonomy and ownership of their research projects to develop increasingly advanced scientific thinking and communication skills. Students undertaking the first iteration of these three vertically integrated courses reported learning gains in course content as well as skills in scientific writing, hypothesis construction, experimental design, data analysis, and interpreting results. Students also demonstrated increasing skills in both hypothesis formulation and communication of findings as a result of participating in the inquiry-based curricula and completing the associated practical assessment tasks. Here, we report the specific aspects of the curricula that students reported as having the greatest impact on their learning and the particular elements of hypothesis formulation and communication of findings that were more challenging for students to master. These findings provide important implications for science educators concerned with designing curricula to promote scientific thinking and communication skills alongside content acquisition.

scientific reasoning; laboratory teaching; written communication skills; student autonomy; student-directed research

WITHIN AN INTERNATIONAL CLIMATE focusing higher education on the task of creating a 21st century workforce who are able to deal with novel, complex, unstructured problems, there is increasing emphasis on the specific pedagogies that will adequately prepare students to meet those expectations. This has caused many science educators to move toward inquiry-based curricula, which are characterized by engaging students in “...identifying questions, attending to evidence, identifying patterns, making controlled comparisons, interpreting increasingly complex data, supporting claims, and drawing justified conclusions” (15). Inquiry-based curricula have been proposed to help students develop these sorts of advanced thinking skills in tertiary science education in general (11) and in physiology education in particular (17). Thus, inquiry curricula have been increasingly used in science education over the last 30 yr to teach students “how to think like scientists” (7).

Inquiry-based curricula have been shown to produce superior learning outcomes compared with traditional curricula (4, 6, 16). However, differences in the contexts and goals of the classroom and research laboratories complicate the translation of the research experience from practicing science as a scientist to the classroom context of practicing science as a student (10). Indeed, the degree of guidance and ownership provided to students during inquiry-based classes has been shown to have a significant impact on the learning outcomes that students achieve (14). Furthermore, students who have become accustomed to, and quite successful in, more traditional curricula often struggle when they first encounter inquiry-based curricula and assessment at the tertiary level (20). Therefore, this study explores the experiences of undergraduate students undertaking a series of vertically integrated practical curricula in biomedical science at the University of Queensland (Brisbane, Queensland, Australia), with a specific focus on the aspects of the design and implementation of these curricula that students perceived as having the largest impact on their learning of disciplinary skills and content.

In 2006, we undertook an extensive review of the curricula across the entire Bachelor of Science (BSc) degree (22). Based on international recommendations for reform in science education (3, 18), we designed a series of vertically integrated courses that used an inquiry-based approach to the practical curricula explicitly aimed to help students develop skills in scientific thinking. Within our context, the key learning goals were for students to develop increasingly advanced skills in scientific thinking and scientific communication by working with increasing ownership of their research questions and autonomy over their research projects. Here, we report an investigation of the specific aspects of the curricula that students reported as having the largest impact on their learning. Our findings provide important implications for science educators who are attempting to develop inquiry-based curricula that promote student learning of a range of disciplinary-specific skills and content.

METHODS

Institutional and Degree Contexts

This study was conducted at a large, research-intensive Australian university and was approved by the institutional ethical review board. The three semester-length (13 wk) courses described in this study are the first- and second-level courses required for students who want to major in biomedical science (specializing in physiology, pharmacology, neuroscience, or molecular and cellular biology). Typically,
students undertake course 1 in the second semester of their first year and courses 2 and 3 in the first and second semesters of their second year, respectively. Course 1 is a recommended prerequisite for course 2, and course 2 is a recommended prerequisite for course 3.

Each course has a designated practical coordinator (faculty member) who oversees the laboratory classes and their associated assessment as part of the teaching load within their academic role. Practical coordinators are responsible for training the teaching assistants, organization of the classes, liaising with technical staff, and teaching into the practical classes. During the period of this study, there were 775, 415, and 380 students enrolled in courses 1, 2, and 3, respectively. Each of the teaching laboratory spaces can accommodate 60–96 students, so each course cohort of students is divided into 6–8 practical groups of 60–96 students. We maintain a teaching staff to student ratio of 1:12, so each practical group has 5–8 teaching assistants, who mentor the same 12 students for the entire semester. Students form self-selected groups of 3–6 students/group in the first practical class in each course and are actively encouraged to remain in these groups, with the same teaching assistant (referred to as “tutors” by the students) for all of the remaining classes. This allows teaching assistants to provide consistent mentoring and feedback to students across multiple assessment items and as the scaffolding is reduced within courses. Several of the teaching assistants provide classes for more than one practical group, so teaching teams consist of 15–22 teaching assistants across the 3 courses; occasionally, teaching assistants also taught into more than one of the three courses, providing an excellent additional source of vertical integration for both the teaching teams and students. Importantly, this demonstrates that the inquiry-based curricula described in this report are scalable to very large cohorts.

Just over half of the students taking these three courses were enrolled in a 3-yr BSc degree or a 4-yr dual degree combining the BSc degree with another degree. Another 20% of the students were enrolled in the Bachelor of Biomedical Science (4-yr research-focused program), and 12% of the students were enrolled in a medical program (6-yr accelerated BSc/Bachelor of Medicine, Bachelor of Surgery program). The remaining students were enrolled in a large range of specialty and applied science degrees. The majority of students undertaking these degrees enter university straight after secondary school, with mid- to high-range university entry scores. By graduation, all of these students are expected to have learned to think and communicate like scientists (22). A review of the BSc degree led to recommendations that “. . . students should have: (i) the opportunity to be critical thinkers who value the opportunity for open-ended inquiry and (ii) who have some insight into how scientific knowledge is generated” (22), and this prompted a move to inquiry-based curricula.

Curricular Context

Overall, course 1, “Cells to Organisms,” sets the scene for students with fundamental concepts around how cells interact to perform complex tissue and organ functions in an organism, with a particular focus on the human body and other higher organisms. Key concepts include basic cellular transport, signaling mechanisms, and an understanding of how structure and function are interrelated in organ systems. Course 2, “Integrative Cell and Tissue Biology,” pulls students deeper into cellular and molecular physiology, with a specific focus on how cells associate and interact to fulfill their normal functions in tissues and organs of the human body. Key topics include epithelial, neuronal, muscle, endocrine, and immune function. Course 3, “Systems Physiology,” builds on this knowledge by extending students’ understanding of human physiology to the level of integrated systems, focusing on integration at the system, organ, cell, and molecular levels. Key knowledge areas include the peripheral and central nervous systems, gastrointestinal tract, renal system and water balance, respiratory system, cardiovascular system, metabolic endocrinology, and glucose homeostasis. All three courses include 13 wk of three 1-h lectures/wk. Course 1 has fortnightly peer-assisted study sessions that have a tutorial-like format, which are optional for students; there are no such tutorials for courses 2 and 3. All three courses have five or six 3-h practical classes throughout the semester, during which students undertake laboratory experiments. Learning goals for the practicals include the ability to design a well-controlled experiment, collect, analyze, and critically evaluate data, integrate this with primary research literature to reach appropriate conclusions, and convey this in oral and written forms of scientific communication.

Design and Implementation of Vertically Integrated, Inquiry-Based Practical Curricula

The model of vertical integration of inquiry-based practical curricula described here has been cited in a nation-wide investigation of the inquiry-based teaching approaches as a representative case study of how the undergraduate curriculum can be redesigned for a focus on inquiry-based learning (8). The design of the vertical integration was guided by the research skills framework developed by Willison and O’Regan (24), which has been successfully adapted for curricula and assessment design in a range of disciplines (23). Specifically, this framework uses a rubric to aid curriculum designers in detailing the specific curricular and learning objectives that are considered to be important in a research-based curriculum across the incremental levels of development that are appropriate to a particular educational context. Our key curricular and learning objectives were for students to develop increasingly advanced skills in scientific thinking and scientific communication and to work with increasing ownership of their research questions and autonomy over their research projects, as shown in Table 1 and detailed below.

Course 1. The learning goals for the practicals in course 1 included becoming aware of the processes that scientists use to discover and verify new knowledge in biology, specifically observation-based discovery and the reproducible testing of explanations through hypothesis-driven inquiry, as well as developing competency in basic biological laboratory techniques.

Practicals consist of five classes of 3 h each, as shown in Table 2. The first class introduces students to the course, where students perform an animal dissection to gain an overview of the systems they will investigate during the semester and to learn the key scientific and technical skills that they will be using in the remaining practicals. Each of the next three classes explicitly require students to design and undertake an experiment to gather evidence that then apply to a simple case study. The final practical class provides students with the opportunity to contrast the animal systems they have been previously investigating with a plant-based investigation.

Students are guided through the practicals by an interactive online laboratory manual built using LabTutor software (AD Instruments), which is used to scaffold students through the basic structure of the scientific article genre. Specifically, the LabTutor system presents students with background information and allows students to enter responses to open-ended questions and choose from multiple-choice questions to develop the hypotheses, methods, results, and discussion sections of the final report. The LabTutor system is also integrated with the PowerLab data-acquisition system (AD Instruments) and Chart software (AD Instruments) to embed physiological data recorded live into student reports. Students submit an individual report 48 h after each practical class; reports are assessed by their assigned practical class teaching assistant, who provides written feedback at least 5 days before the next practical class.

Course 2. Learning goals for the practicals in course 2 include extending students’ skills in developing research questions, designing and conducting experiments, and developing skills in statistically analyzing data, interpreting results, and writing scientific proposals and reports.

Practicals consist of two projects, each consisting of three 3-h classes (Table 2). During the first project, students design and conduct...
Table 1. Vertical integration of the curricular objectives (student autonomy and ownership) and learning objectives (skill development) across the three inquiry-based practical curricula

<table>
<thead>
<tr>
<th>Course code and name</th>
<th>Course 1</th>
<th>Course 2</th>
<th>Course 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIOL1040: Cells to Organisms</td>
<td>BIOM2011: Integrative Cell and Tissue Biology</td>
<td>BIOM2012: Systems Physiology</td>
<td></td>
</tr>
<tr>
<td>Year 1, semester 2</td>
<td>Year 2, semester 1</td>
<td>Year 2, semester 2</td>
<td></td>
</tr>
<tr>
<td>Six 3-h classes</td>
<td>Two blocks of three 3-h classes</td>
<td>One block of five 3-h classes</td>
<td></td>
</tr>
<tr>
<td><strong>Curricular objectives</strong></td>
<td><strong>Vertical Progression</strong></td>
<td><strong>Skill-building class</strong></td>
<td><strong>Skill-building experiments</strong></td>
</tr>
<tr>
<td><strong>Autonomy</strong></td>
<td>Increase</td>
<td>LabTutor provides specific scaffolding for each stage of the experiment from hypothesis to discussion. Teaching assistants provide progressively less guidance to students across the semester.</td>
<td>Students are given freedom to investigate any aspect of cardiovascular, respiratory, renal, and metabolic physiology in response to a wide range of perturbations.</td>
</tr>
<tr>
<td><strong>Scientific process</strong></td>
<td>Increase in complexity</td>
<td>Hypothesis generation and experimental design are emphasized. Strategic questions scaffold the entire report writing process.</td>
<td>Students are expected to be working autonomously for the second half of the semester.</td>
</tr>
<tr>
<td><strong>Scientific communication</strong></td>
<td>Reduction of scaffolding and more advanced aspects of scientific writing.</td>
<td>Formal proposals and formal reports are completed. Students are provided guidelines and lectorials. The emphasis is on genre conventions for methods and results.</td>
<td>Students are expected to be working autonomously for the second half of the semester.</td>
</tr>
<tr>
<td><strong>Assessment</strong></td>
<td>Increase in complexity</td>
<td>Three reports</td>
<td>Two experimental plans, two proposals, and two reports (1 report/module)</td>
</tr>
<tr>
<td><strong>Project duration</strong></td>
<td>Increase</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Timing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Students typically undertake course 1 in the second semester of their first year and courses 2 and 3 in the first and second semesters of their second year, respectively.

experiments on double-pithed toads with excised hearts to investigate the impact of one or two factors from a large range of options on cardiovascular parameters. For the second practical project, the majority of students investigate skeletal muscle physiology, where students are the subjects for their own experiments using electromyography (EMG) recordings to examine the impact of one or two factors, from a long list of options, on the relationships between electrical and mechanical activity of human skeletal muscles. Instead of the EMG experiments, a small subset of students who were undertaking the research-focused Bachelor of Biomedical Science investigated the regulation of macropinocytosis using a cell culture paradigm (5).

As shown in Table 2, each project module begins with a 3-h skill-building class where students gain hands-on experience with the experimental paradigm and develop their experimental design. Students are then provided with a 1-h interactive lecture (lectorial) with the practical coordinator and a 1.5-h practical class with their teaching assistants. For the first module, these short classes are used to guide students through the data-analysis conventions and procedures using their pilot data from the first class. In the second module, the 1-h lectorial and 1.5-h practical class are used to guide students through the conventions and expectations for writing scientific papers using feedback on their marked reports from the first module. Each project module then ends with a 3-h class in which students perform the experiment they have designed and collect the data.

During each module, students are assessed on the following three written tasks: 1) their experimental design developed as a group during the first class of each module; 2) a group proposal including an introduction, hypothesis, and refined methods section completed before the final class of each module; and 3) an individual report in the form of a scientific journal article due 1 wk after the final class of each module. Each of the three assessment items within a project module are linked, so that feedback on each assessment is directly relevant to subsequent assessments.

Course 3. In the course 3 practicals, students design and conduct a larger-scale semester-long clinical experiment to achieve learning goals of collaborating on a large scale, developing more complex experimental designs, and thus conducting more advanced experiments, acquiring and analyzing larger sets of quantitative data as well as interpreting their findings in relation to published literature at a broader and more indepth level than in course 2.

As shown in Table 2, all five 3-h practical classes in course 3 contribute to one major project, designed and conducted by the students,
How We Teach

INQUIRY PRACTICALS DEVELOP SCIENTIFIC THINKING SKILLS

Table 2. Summary of the organization of the practical component of the three consecutive, vertically integrated courses

<table>
<thead>
<tr>
<th>Course 1</th>
<th>Course 2</th>
<th>Course 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practical 1</td>
<td>Practical 2</td>
<td>Practical 3</td>
</tr>
<tr>
<td>Toad dissection experiment</td>
<td>Osmosis experiment</td>
<td>Action potential experiment</td>
</tr>
<tr>
<td>Pilot experiments on toad heart (skill building and experimental planning) (module 1)</td>
<td>Data analysis lectorial and computer class (module 1)</td>
<td>Experimental data collection (module 1)</td>
</tr>
<tr>
<td>Structured experiments on blood flow and volume (skill building)</td>
<td>Structured experiments on oxygen and energy use during exercise (skill building)</td>
<td>Oral presentation of student experimental proposals</td>
</tr>
</tbody>
</table>

Students undertake five separate 3-h practicals in course 1, there are two modules of three 3-h classes in course 2, and there is one extended series that takes place across five 3-h classes in course 3. Based on student feedback, the final course was later modified to include a sixth class on data analysis. N/A, not applicable.

Investigating the integrated functions of several physiological systems involved in restoring homeostasis in response to a perturbation (usually exercise). The project begins with two skill-building classes where students gain hands-on experience with the equipment available for use in their project. After completion of the second skill-building class, each research group collaboratively develops a research question and drafts a hypothesis and the methodology to examine their research question. During the third practical class, each group formally presents their research proposal to their peers in an oral presentation. At the end of this class, one of the presented proposals is selected by the students by a majority vote for the whole practical group in a class (60–72 students) to perform for the remaining two practical classes. After the vote, a whole class discussion of specific details of the experimental design is led by the teaching assistants, so that the experimental design can be improved when needed. Students conduct the chosen experiment in the final two practical classes and pool the data from the whole class cohort, and each student then independently analyzes the results and submits a written report for the project.

Students are assessed on their draft hypotheses and methods, group research proposals and presentations, and individual final report. Support from teaching staff consists largely of advice and feedback on the experimental design during the second and third classes and suggestions for the analysis of results and presentation of findings as data are collected in the final two classes. Overall, this means that in course 3, students are much more reliant on peer support and making their own decisions throughout the entire extended project than they were in the previous two courses.

Specific Scaffolding for Skills in Scientific Thinking and Communication

As shown in Table 1 and described in detail above, students work with increasing ownership of their research questions and autonomy over their research projects to achieve the learning goals in developing scientific thinking and communication skills as they progress throughout the three courses. Two specific outcomes of these learning goals are the generation of appropriate scientific hypotheses and the appropriate analysis and representation of experimental results. Therefore, in this study, we analyzed the quality of students’ hypotheses and figure legends from reports submitted for all three courses, and the vertical integration of the scaffolding of these specific aspects is described in more detail below.

Hypothesis scaffolding. In the first practical of course 1, students were provided with three hypotheses and asked to choose the most appropriate one. For the subsequent three reports, students were required to write their own hypotheses and were provided with feedback by teaching assistants in class. In course 2, students were expected to develop their own hypotheses and were provided with feedback in each of the three assessment submissions: the experimental plan, the proposal, and the reports. In course 3, this scaffolding and guidance were again reduced, with students developing their experimental proposals outside of class without tutor feedback; the hypothesis was also finalized by students independently after the completion of their final experimental class.

Figure legend scaffolding. For the first report in course 1, figure legends were not used. For the second and third reports, although figure legends were not emphasized by teaching assistants or marked in the reports, the LabTutor system did provide students with example figure legends, which consisted of two statements. The first statement summarized the scope of the figure, whereas the second statement provided a description of the data presentation (Fig. 1A).

In course 2, students had two classes in which they were given explicit instructions and feedback on how to write figure legends. Again, an example figure legend was provided, which included a description of the data presentation and relevant experimental conditions followed by a statement describing the statistical information shown in the figure (Fig. 1B).

In course 3, all scaffolding was then removed, with students completing the data analysis and figure legends for their reports after their final practical class without tutor guidance. Instead of being provided with specific examples of figure legends, in course 3, students were prompted to go to the literature they were using for their assignments to find examples to guide their figure legend writing.

Data Collection and Analysis

This study has taken both a snapshot and longitudinal approach to the evaluation of the vertically integrated curricula. First, we analyzed the course evaluations provided by the first cohort of students who undertook each of these three courses to determine which aspects of the curricular design had the greatest impact on student learning for each of the four learning objectives shown in Table 1. Second, we analyzed a sample of 21 students’ reports from all 3 courses to investigate the learning gains in scientific thinking and communication achieved by students as they progressed through the vertically integrated inquiry-based practical curricula.
Data from course evaluations. Course evaluations were collected as follows. For course 1, students were asked a series of reflective questions at the end of their final practical assessment submission. For this report, we analyzed the responses to the following questions: “Which practical class was most helpful for your learning and why?” and “Which practical class was least helpful for your learning and why?” For courses 2 and 3, extensive online course evaluations were conducted at the end of each semester through the institutional teaching evaluations unit, and students were provided with a small amount of course credit for completing the surveys (~1% of the course grade). The responses to the following questions were included in this analysis of the practical curricula: for course 2, “Please comment on how the COURSE ACTIVITIES helped your learning” (emphasis in original), and for both courses 2 and 3, “Please comment on how the ASSIGNMENTS AND ASSESSMENTS helped your learning” (emphasis in original), “What do you think are the strengths of this course,” and “What improvements would you suggest in this course?”

Course evaluations were provided by only 149 of the 775 students (19%) enrolled in course 1. In contrast, providing students with the small reward of ~1% of the course grade elicited dramatic increases in the response rates for courses 2 and 3, where 395 of the 415 students enrolled (95%) and 306 of the 380 students enrolled (81%) completed the evaluations, respectively. Of this corpus of course evaluations, comments about the practical curricula were provided by 140 students for course 1 (94% of respondents, 18% of the students enrolled), 254 students for course 2 (64% of respondents, 61% of the students enrolled), and 220 students for course 3 (72% of respondents, 58% of the students enrolled). All evaluations were deidentified before further analysis so we were unable to match responses from individual students across the three courses and have thus provided snapshots of how the cohorts as a whole evaluated the inaugural iterations of these curricula.

K. Zimbardi applied three levels of inductive thematic analysis to the course evaluation data following the step-by-step procedure detailed by Hatch (12). Briefly, this involved reading and rereading the entire data set, developing themes from the data to address the research questions (which were “What do students report learning from the practical curricula and how?”), comparing the themes to develop hierarchical trees of related themes, looking for data that contradicted the themes and revising and refining the themes based on this, comparing subcategories within themes to further understand the meaning of the themes, and comparing between themes to refine the relationships between themes. In the first level of inductive analysis, three major themes were found for what students reported learning from the practical curricula (learning in general, learning content knowledge, and learning skills). In the second level of inductive analysis, six themes emerged from the data for how students learned content and skills: 1) assessment (how completing the assessment tasks impacted on learning), 2) hands-on experience (how gaining practical experience impacted learning), 3) relatedness (how practical task were related across practicals and to lectures), 4) scaffolding (how guidance and scaffolds from the teaching assistants, practical manuals, assessment criteria, etc. impacted learning), 5) autonomy (how having ownership over the research questions and autonomy in conducting the research projects impacted on learning), and 6) difficulty (how the degree of difficulty or ease of tasks impacted on learning). Finally, in the third level of inductive analysis, comments were classified as positive if students reported that the aspect of the curricula had helped them to learn content knowledge or a particular skill or negative if students reported that their learning of that aspect was hindered by an aspect of the curriculum or could have been greater if the curriculum was changed.

Data from student scientific reports. At the end of the second semester in 2009, we randomly selected two classes of students in course 3 and e-mailed each of the students to ask for informed consent to use their assessment items for research purposes. Of the 27 students who agreed, 21 students had completed the inaugural offering of courses 1–3. These 21 students were representative of the cohorts for each of the courses; Mann-Whitney U-tests for categorical and ordinal data were used to determine that there were no significant differences (P > 0.05) between the final course grades of this sample of students and the cohort for courses 1, 2, and 3 or significant differences (P > 0.05) based on sex or on whether they were domestic or international students. For the inaugural offering of course 2, completion of the second module and the associated assessment were not compulsory. As only 5 of these 21 students submitted the second report for course 2, these second reports from course 2 were excluded from the following analysis. Therefore, all three reports from course 1, the report from the first module in course 2, and the report from course 3 were used for the following qualitative analysis.

The hypotheses and figure legends from each report were imported into NVivo (QSR) and analysed by K. Zimbardi following the step-by-step procedure for typological analysis of qualitative data described by Hatch (12). Briefly, each of the hypotheses and figure legends was inspected looking for the types of information and statements that the authors, as researchers in physiology and pharmacology, expected to be included scientific hypotheses and figure legends (namely, dependent and independent variables for hypotheses as well as data presentation and statistical information for figure legends). Across the entire data set, four different elements were found in the hypotheses and five elements in the figure legends that agreed with, and extended, the typology of statements the authors had predicted would be present in the data. Each hypothesis was then inspected for the presence or absence of each of the four elements: 1) dependent variables, 2) independent variables with 3) comparative statements, and 4) context (see examples from student reports in Fig. 2). Each figure legend was inspected for the presence or absence...
of each of the five elements: 1) scope, 2) results, 3) data presentation, 4) statistical significance, and 5) experimental details (see an example from a student report in Fig. 3A). We also applied the same analysis for figure legends from bar and line graphs in a sample of 10 published articles cited by the students (see example from the Journal of Physiology in Fig. 3B). A range of journals was included in the analysis from journals publishing articles of broad significance in physiology (for example, the Journal of Physiology and European Journal of Applied Physiology) to journals focused on publishing articles in specialist areas of physiology (for example, Hypertension and the American Journal of Cardiology) with impact factors ranging from 7.45 to 1.63.

The presence or absence of each element in each hypothesis and figure legend was determined for each of the 5 reports for each of the 21 students. As this represented binary longitudinal data, binary logistic regressions were performed (SPSS version 21, IBM) to determine whether there were statistically significant increases in the inclusion of each element across the successive reports for sample of 21 students. P values of <0.05 were considered statistically significant, and Nagelkerke $R^2$ values are reported.

### RESULTS

#### Course Evaluations

Students referred to learning in the general sense (186 comments), learning content knowledge (503 comments), and learning skills (232 comments). Comments about learning in the general sense did not specifically indicate what the students were expecting to learn [e.g., “The strengths of this course is the prac as they are very useful to help students learn” (course 3)] and were not analyzed further.

#### Learning Content

Student comments on learning content knowledge related to the lectures and/or exams or specific concepts (such as osmosis, action potential propagation, and cardiovascular function) were analyzed to determine which aspects of the practical curricula impacted most on content learning (Fig. 4). This revealed a striking change between the first- and second-year courses in the...
features of the curricula that students most frequently reported as helping their content learning. In course 1, the most frequent positive and negative comments were focused on how well the practicals were scheduled to align with the concurrent lecture concepts (Fig. 4A, relatedness). This was followed by comments about gaining hands-on experience in practicals (Fig. 4A, hands-on) that students said allowed them to “apply” their content knowledge in a practical setting (either explicitly: “Practical allows me to apply what I have learned in lectures and further my understanding” (course 1) or implicitly “I learn well putting knowledge into action” (course 1)) or being able to “see” the concepts in action (either explicitly: “it really helped me see the effect on a physical level so that I could understand the concept in more depth and not just as superficial, rope-learnt [sic] information” (course 1) or implicitly “Skeletal muscles [practical was most helpful to my learning] because the very twitching of the muscle helped burn the idea of contraction into my mind” (course 1)). In contrast, for both of the second-year courses, the most common comments indicated that completing the written report after each project was pivotal in assisting students with their learning of content as a proportion (%) of all student comments about learning content for each course. A: course 1. B: course 2. C: course 3.

**Learning skills.** The majority of the comments on learning technical skills were provided for course 1 (Fig. 5). The most common skill students mentioned involved learning to dissect the experimental animal in courses 1 and 2 or to take clinical measurements in course 3. The majority of the positive comments indicated that simply gaining experience in performing these techniques made the greatest impact on their learning (Fig. 6, A–C: hands-on). Positive comments about the impact of the level of difficulty on earning technical skills (Fig. 6, A–C: difficulty) indicated that students found the technical tasks challenging, whereas the negative comments indicated that students were not being challenged enough (tasks were too easy). Similarly, positive comments about the relatedness (Fig. 6, A–C: relatedness) indicated that the technical skills were new to them and that overlaps between practicals helped to cement their learning of these skills, whereas negative comments indicated that technical skill learning was hindered when the skills were not new to them or quickly became monotonous. Overall, this would suggest that learning technical skills is a preliminary part of the learning in practicals that quickly gives way to more advanced learning once the technical skills are mastered.

**Learning skills in scientific communication.** The majority of comments students made about learning skills in scientific communication were provided for course 1 (Fig. 5). The most common skill students mentioned involved learning to write laboratory reports, which were often viewed as an opportunity to practice writing skills. Positive comments about the relatedness (Fig. 6, A–C: relatedness) indicated that the written reports helped students to see the concepts in action (either explicitly: “it really helped me see the effect on a physical level” (course 1)) or being able to “see” the concepts in action (either explicitly: “I learnt a lot about our groups [sic] particular topic chosen for the practical report, but while this is interesting it may not have a huge amount of relevance for this course. I would rather that we have a set practical to do that is going to help us more with our understanding of the course material.” (Course 3))

Thus, alignment between the practicals, associated assessment, and lectures is seen by students as one of the most important aspects of how practicals support their content learning and, importantly, is related to the degree of autonomy students have in designing their experiments.

**Learning Skills**

The comments students made about learning skills described the following three broad categories of skills: technical skills, scientific communication skills, and skills in the scientific process (Fig. 5).

Fig. 4. How students learn content.Shown are each of the aspects of curricular design that students said helped (positive comments) or hindered (negative comments) their learning of content as a proportion (%) of all student comments about learning content for each course. A: course 1. B: course 2. C: course 3.

Fig. 5. Types of skills students reported learning during the practical curricula and associated assessment. The numbers of comments that students made in the course evaluations after courses 1, 2, and 3 regarding their learning of technical, communication, and scientific thinking skills during the practical curricula and associated assessment are shown.
communication came after course 2, as shown in Fig. 5. The vast majority of the comments students made about learning communication skills related to skills in scientific writing and indicated that gaining experience in writing “proper” scientific reports that were modeled on scientific articles for their assessment helped students develop their skills in scientific communication (Fig. 6, D–F: assessment), for example:

Since I was totally unaware of the intricacies of scientific report writing, having two opportunities to write one meant I could waste the first one discovering this fact, and the second figuring out what...I was doing.

(Course 2)

Students also indicated that having multiple written assessment tasks (Fig. 6, D–F: assessment) that built on each other (especially the experimental plan, proposal, and report in course 2) with feedback from teaching assistants (Fig. 6, D–F: scaffolding) was important for helping students to develop skills in scientific writing, as follows:

The proposal was a great because we got to hand in a ‘draft’ report and the feedback my tutor gave me really helped me to improve and write a better report. I think the feedback I got on both the proposal and the report will help me to write reports next semester.

(Course 2)

Importantly, students also explicitly indicated that the dual purpose of using report writing opportunities to learn skills in scientific communication and content needed to be balanced appropriately for each assessment task, for example:
It was very useful though having such an easy prac first up because it allowed us to learn how to write up these pracs properly and given that the topic was easy, it allowed us to focus more on our style of writing than on having to fully understand the concepts as well (if we didn’t already).

(Course 1) Although the oral presentation task in course 3 was designed to help students develop skills in scientific presentations, particularly in developing persuasive verbal arguments based on scientific literature, there were few comments from course 3 about learning skills in oral presentation. Most comments suggested that the marking standards may have been too easy (Fig. 6F: difficulty) and thus did not challenge some students to develop their skills beyond their current capabilities [e.g., “the presentation was easy marks but pointless” (course 3)].

Learning scientific thinking skills. The majority of comments made about learning skills associated with the scientific process attributed specific learning gains in hypothesis construction, experimental design, interpreting results, and data analysis to having autonomy in developing their own experiment (Fig. 6, G–I: autonomy), for example:

The osmosis prac class is the one I found most useful. It was during this prac that we first had to develop a hypothesis and method of our own to test it.

(Course 1) …the practicals helped to teach proper experimental design at a level not done before in another course.

(Course 2) [Strengths of the course were] designing our own prac to understand the process of how this is done.

(Course 3) In addition, having to write about the basis for the design of their own experiment, and the analysis and interpretation of their own results in their assessment was influential in helping students to learn these skills in scientific process (Fig. 6, G–I: assessment), for example:

The practical were the most helpful for me. I learnt so much from having to write a detailed scientific paper on my findings and my research design.

(Course 2) The presentation and talk really helped my learning as I found it good to understand the step by step process necessary for preparing an experiment.

(Course 3) The impact of having to design their own experiments was also found to cause students to undertake more complex data analysis and interpret more complex results than they had experienced in other more structured courses, for example:

Designing our own experiments were good though, and the statistics was even more involved than in my statistics course!

(Course 3) However, in course 3, students also indicated that there was too much autonomy for the more complex aspects of the data analysis and interpreting results (Fig. 6f: autonomy). Therefore, in many cases, students also specifically requested additional guidance:

There was no assistance with Prism [statistical and graphing software] except from tutors. But practical classes had finished when we were completing the assignment so we could no longer discuss in person the assignment with tutors.

(Course 3) Students also requested more guidance from the instructors or more structure to the discussions to improve the experimental design, for example:

Vote for final experiment differently and have a chance to put forward ideas for changes.

(Course 3) Good setup making students design their project, however I would’ve liked more time to discuss the physiology of the chosen groups design.

(Course 3) Student Scientific Reports

Analysis of hypotheses. In the hypotheses, dependent variables were always present (Fig. 7); however, some of the dependent variables in the reports from course 1 were not measurable by the methods available to the students, for example:

We hypothesise that cells in a hypotonic solution will explode while cells in a hypertonic solution will shrink down. During their experiments, students were not able to directly measure cells exploding or shrinking (although this represents a fair conceptual understanding); instead, they were able to measure the absorbance of the solution (indicated by the color of the solution) using a spectrophotometer. Therefore, an example of a student hypothesis with a measurable outcome in this context was the following:

A hypotonic solution will result in a greater amount of colour in the blood plasma, in comparison to a hypertonic solution.

There was a statistically significant increase in the number of students in our sample who included measurable dependent
variables in their hypotheses across the five reports ($P < 0.05$, $R^2 = 0.482$; Fig. 7). Students always included independent variables that were appropriate to the methods available (Fig. 7) but did not always include explicit comparisons for the independent variables, generally omitting the comparisons with a baseline control, as illustrated by the following example:

Anticipation of exercise will cause an increase in Heart Rate and Blood Glucose levels compared with hypotheses that explicitly included comparisons with control or baseline conditions:

The addition of adrenaline to a propranolol treated toad heart will not produce any observable changes in heart rate, compared to a baseline reading

or

Anticipation of exercise would result in increased blood glucose in the period prior to exercise compared to individuals not anticipating exercise.

There was a statistically significant increase in the number of students in our sample who included comparator statements for the independent variables in their hypotheses across the five reports ($P < 0.05$, $R^2 = 0.325$; Fig. 7).

In contrast, although the hypotheses usually included contextual information about the experiment (i.e., the animal and/or tissue on which the experiment was being performed), the presence of this information did not change significantly across the five reports ($P > 0.05$; Fig. 7).

Analysis of figure legends. Across all four assignments for which students wrote figure legends (reports 2–5), there was almost always a statement that summarized the scope of the figure. In many cases, these scoping statements matched the independent and dependent variables and context used in the hypothesis for that report, for example:

The effect of a local anesthetich in the compound action potential amplitude in the sciatic nerve of a Bufo marinus over time.

There was a statistically significant increase in the number of students in our sample who included result descriptions ($P < 0.05$, $R^2 = 0.147$; Fig. 8), data presentation ($P < 0.05$, $R^2 = 0.239$; Fig. 8), statistical information ($P < 0.05$, $R^2 = 0.453$; Fig. 8), and experimental information ($P < 0.05$, $R^2 = 0.171$; Fig. 8) in the figure legends across reports 2–5 but no change in the use of scope statements ($P < 0.05$; Fig. 8).

Overall, the information that students included in their figure legends across the three courses is consistent with the information found in figure legends in published literature. There was a large degree of variation in the types of information included among the small sample of figure legends from published articles (Fig. 8). However, it is important to note that, first, students did not include types of information not found in published figure legends and, second, published figure legends did not include any additional types of information not found in the majority of students’ figure legends in courses 2 and 3.

DISCUSSION

The series of vertically integrated inquiry-based practical curricula described in this report was designed to help students develop the critical thinking skills needed to solve the complex problems they will face after graduation. Specifically, these curricula aimed to facilitate student development of increasingly advanced skills in the scientific method and scientific communication by working with increasing ownership of their research questions and autonomy over their research projects. This study found that students reported substantial gains in their learning of content and skills in laboratory techniques, scientific communication, and scientific thinking. Learning gains in scientific communication and thinking were evident in the hypotheses and figure legends written by students in their five scientific reports as they progressed through the three courses. Key features of the curricular design and areas for improving the design to further support student learning were identified from student evaluations and the nature of the changes in student hypotheses and figure legends.

Providing students with multiple opportunities to gain experience and feedback in developing their own hypotheses and experimental designs clearly helped students become familiar with the process and conventions for scientific investigation and to become acutely aware of the central role that hypotheses and experimental design play in biomedical science. This is particularly important given recent evidence showing that conceptions of hypotheses represent a threshold concept in biology (21), that is, students find the rules for effective hypothesis construction as well as the roles that hypotheses play in science particularly troublesome concepts to understand, but once they do begin to understand these important scientific principles, their thinking is irreversibly transformed. Developing student understanding of hypotheses has been found to be the most commonly cited learning goal for inquiry-based practicals in undergraduate biomedical science (8). This is certainly appropriate considering the central role that hypothetico-deductive reasoning has been show to play in scientific thinking (7) and in sound reasoning more generally (15). Indeed, we (25) have recently found that students entering course 1 carry a range of conceptions around hypotheses. These conceptions range from a simplistic view that hypotheses are based on facts to more advanced and accurate conceptions that hypotheses are predic-
tions and multivariate causal relationships that need to be tested using evidence from controlled experiments. The curricula we have described in this report serve as key examples of how the basic premises underlying hypothesis construction can be embedded in the first year of an undergraduate degree and then built on with increasing complexity through links to experimental design, statistical analysis, and interpretation of results in subsequent courses. In this study, we have shown that students often begin course 1 by writing hypotheses with dependent variables that stem from a conceptual understanding of the physiological mechanisms underlying their experiment rather than the measureable dependent variables that are then linked to the development of suitable experimental methods. In addition, although students recognized the need for independent variables in their hypotheses and chose relevant independent variables, they often missed part of the comparison statement for the independent variable. This is particularly interesting given that one of the key differentiators found between undergraduate students who are nonscience majors, undergraduate students undertaking a biology major, and biological scientists is the extent to which the group identified baseline controls as a necessary control for testing hypotheses (1). Therefore, incorporating baselines or nontreatment controls as part of the independent variable condition represents an important learning goal in developing hypothetic-deductive reasoning in biological science and needs particular attention in curricular design, implementation, and assessment.

Another aspect of hypotheses that students seemed to struggle with was the context of the experimental paradigm. The physiological context is important for determining the limitations to the conclusions that can be drawn from the study and is thus very important for the more advanced levels of scientific reasoning necessary when students discuss the findings of their experiments in relation to scientific literature. For example, there is a large difference between the physiological mechanisms in play for a group of young healthy humans compared with older humans or elite athletes. We have used this specific finding to improve subsequent iterations of this curricula, that is, we have included the specific wording of “experimental context” in the marking criteria for hypotheses in the reports in later iterations of courses 2 and 3 to help students realize the necessity of this information. In the future, we plan to investigate the impact of this strategy on students’ use of context in their hypotheses and in the critical integration of their findings with primary research literature in the discussion sections of their reports.

There was a large degree of variation in the types of information included in the sample of 10 professional figure legends, which raises the question of what information is essential for a figure legend in science. Apart from the well-known “rule of thumb” that figure legends should contain all of the information necessary to understand the figure without reference to the text of the article, there is little in the scientific literature or educational literature on the essential features of figure legends. All 10 figure legends from physiology journals we analyzed did include a scope statement and, for figures with quantitative data, a description of how these data are presented. On the other hand, the additions of experimental and statistical information as well as the results shown in the figure appear to be more context dependent. Regardless of this variability in the scientific literature, using the learning scaffold described in the present study, students’ figure legends increased in complexity in courses 2 and 3, incorporating more of the types of information found in the professional figure legends published in physiology journals.

Findings from student evaluations also demonstrated that students required more explicit instruction than they were provided to develop their skills in statistical analysis, particularly as the analyses became more complex in the large clinical studies in the most senior course. In subsequent iterations of these curricula, we developed courses 2 and 3 to improve the vertical integration of skills around statistical analyses. In course 2, where the experiments are relatively less complex, we provided students with more freedom in developing hypotheses that involve more complex relationships and then scaffolded students in gaining experience in dealing with the complex statistical analyses required to test these hypotheses. In course 3, we then embedded more explicit instruction on the links between experimental design and statistical analysis early in the semester, and we included an additional final practical class that focuses on statistical analyses and interpretation of experimental findings. Not only does this increase student preparedness for the complexity of the analysis undertaken in course 3, but teaching assistants are now also present when students are completing the statistical analysis and are thus able to guide students in making sound decisions about the analysis they undertake and in interpreting the findings. Together, these changes have increased the degree of vertical integration across courses, enabling students to tackle complex analytical techniques successfully and make appropriate interpretations about the results.

Writing reports in the genre of scientific articles was cited as a crucial aspect in helping students to develop their skills in scientific communication. Across both the positive and negative comments, students indicated that learning to write in the scientific genre was an important, advanced skill that required practice and guidance through regular feedback. The aspects of the practical curricula design that students identified as providing the greatest support developing their skills in scientific communication were the authenticity of the assessment design, along with the scaffolding and guides, and the repetitive opportunities for students to develop their written communication skills through experience and feedback from teaching assistants, within each course and across the vertically integrated courses. These aspects all combined into a powerful system for facilitating student learning of effective written scientific communication.

In contrast, students very rarely reported learning gains in oral scientific communication skills, highlighting important flaws in the design of the oral assessment. Specifically, students in course 3 were dissatisfied with the way that the projects and experimental designs were decided during the whole class discussion. In addition, some comments also suggested that high levels of expectations used for the reports need to be applied to the oral presentations to help students develop these communication skills adequately. Thus, an increase in expectations was required to promote greater learning gains, but it was also clear that we could not expect students to directly transfer the experience and skills they had gained in the small group discussion around hypotheses and experimental designs in courses 1 and 2 to the large group oral critiques in course 3. In subsequent iterations of course 3, we provided...
How We Teach

Creating a vertically integrated series of courses in which the early practical classes are extensively guided and scaffolded and then progressively reducing that scaffolding as students progress encourages students to work with increasing ownership of their research questions and autonomy over their research projects. However, the findings of this study also highlight that to appropriately support students when novel types of skills or a large increment in the advancement of the skills is required, there is a need to reintroduce key aspects of the scaffolding temporarily. We know that students struggle when required to take on a level of autonomy and ownership they had not previously encountered (14, 20) but that this autonomy is necessary for graduates to succeed in the workforce. Therefore, the curricula described here serve as important examples of how to provide early stage undergraduate students with appropriately adjusted experiences and levels of autonomy that improve their scientific communication and scientific thinking skills and thus prepare them to confidently and effectively confront the complex, novel problems they will encounter in the 21st century workplace.

ACKNOWLEDGMENTS

The authors thank all of the students who provided the detailed evaluations of these courses that enabled us to determine the specific aspects of the curricula that had the greatest impact on student learning. The authors are also grateful to the Faculty of Science, The University of Queensland, for supporting the development of the evaluation tools and for providing writing retreats and workshops to complete the analysis and writing of this report.

GRANTS

A University of Queensland Teaching and Learning Strategic Grant supported the characterization of the inquiry-based practical curricula.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS


REFERENCES