We set out to develop an anatomy and physiology (A&P) curriculum with a content that was relevant and well rationalized. To do this, we developed a benchmarks curriculum process that helps us to determine what our students need to learn in A&P and then to make sure there is alignment between those learning objectives and what we teach and how we assess our students. Using the benchmarks process, we first set the broad skill and content goals of the course. We then prioritize the topic areas to be covered, allocating the course time accordingly, and declare the learning objectives for each topic. To clarify each learning objective, a set of benchmark statements are written that specify in operational terms what is required to demonstrate mastery. After the benchmarks are written, we assemble the learning activities that help students achieve them and write assessment items to evaluate achievement. We have implemented the curriculum in a relational database that allows us to specify the numerous links that exist between its different elements. In the future, the benchmarks model will be used for ongoing A&P curriculum development with geographically distributed contributors accessing it via the World Wide Web. This mechanism will allow for the continuing evolution of the A&P curriculum.

Benchmark: a mark on a permanent object indicating elevation and serving as a reference in topographical surveys and tidal observations

At some point in a teaching career, a startling realization may arrive—what students learn is sometimes quite different from what we think we are teaching them. Often less obvious is the possibility that what we set out to teach may be quite different from what students need to learn (1). We were faced with both of these possibilities when six years ago we began teaching human anatomy and physiology (A&P) to undergraduates. Fortunately, at the outset neither of us had any prior experience with such a course, having previously only taught pharmacology to medical students and animal behavior to undergraduates, and only knew the existing syllabus was ancient and in need of revision. Not knowing exactly what we were supposed to teach, we were compelled to ask ourselves and others what a contemporary A&P course should include. A&P texts, of which there are many, all cover the usual encyclopedic list of topics, often from a largely anatomic perspective, and did not give useful guidance about course structure. The Human Anatomy and Physiology Society (HAPS) has a curriculum guide (6), but it is also a comprehensive list of A&P topics and does not specify which are most important and what level of learning is expected. APS does not have undergraduate curricular guidelines for A&P.
Steps

1. **Define Broad Course and Curriculum Goals**
   Examples: Develop problem-solving ability and predictive level of understanding

2. **Establish High, Middle and Low Priority Areas**
   Examples: High priority - nervous, cardiovascular system
               Low priority - aging, integumentary system

3. **Establish Objectives for Each Unit (musculoskeletal, cardiac, nervous system etc.)**
   Example: Understand the mechanical activity of the heart and the relationship between rate, filling (EDV) and emptying (ESV, EF)

4. **Establish Benchmarks which Define Operationally How to Know When an Objective is Reached**
   Example: Predict CO based on variation of HR, SV, EDV and ESV

5a. **Assemble Learning Activities**
    - Text Reading
    - Hands-on lab exercises
    - Computer Simulations
    - Computer Animations
    - Workshop exercises
    - Class-time (lecture)
    - Computer Aided Learning (CAL)

5b. **Write Assessment**
    - Formative:
      - Questions in lab exercise
      - Workshop problems
    - Summative:
      - Quizzes
      - Practical
      - Exams

**FIG. 1.** Developmental steps of the benchmarks curriculum model. Arrows and numbers indicate the order in which steps are carried out. EDV, end-diastolic volume; ESV, end-systolic volume; EF, ejection fraction; CO, cardiac output; HR, heart rate; SV, stroke volume.

Not knowing the topography of the A&P terrain, we began with a list of questions to help us chart our way.

1) How do we identify what students should learn?
2) How do we decide topic priorities?
3) How do we create course content and pedagogy that is aligned with our priorities and learning objectives?
4) How do we communicate the learning objectives to students?

5) How do we assess student success in achieving the learning objectives?

After six years of attempting to answer these questions, the result is a new and evolving A&P course and, perhaps more important, a new process for developing a course or curriculum. This “benchmarks curriculum model” is diagrammed in Fig. 1 (9). The reasons for the name will be clearer as the process is described. Suppressing the impulse to immediately begin assembling topics, syllabi, lab exercises, textbooks, and other course materials is not easy, but the benchmarks process has helped us establish a sequen-
tial approach that places this activity after the careful definition of what the course is meant to accomplish.

THE BENCHMARKS PROCESS

How do we identify what students should learn?

What are the course goals? First, we had to find out what students needed to have when they finish the course. This is step 1 of Fig. 1. Unlike the medical school curriculum, in which the outcomes are perhaps better defined, the course goals for A&P are less obvious. There is not one set of skills and knowledge that all A&P students need, because A&P is a gateway to many allied health and other fields of study. We needed to identify a common subset to be covered that prepares students well for a variety of career paths. Identifying the crucial set of course goals led us into consultation with instructors of courses that follow A&P and a review of some of the clinical licensing exams (15). Both helped define the content coverage and the level of understanding expected of successful students of A&P. Because most of our students planned nursing majors, we consulted extensively with our nursing school and reviewed the practice licensing exams for nursing. This process was enlightening and immediately highlighted the misalignment between the existing course and the actual expectations about our students’ preparation for subsequent courses. A&P is traditionally memory intensive with regard to structure, and we learned that much of this information was not needed either in subsequent clinical courses or licensing exams. On the other hand, we learned that our students needed a much better functional understanding of, among other things, nervous system function and fluid balance.

The type of learning the nursing faculty was looking for is exemplified in the question, “During severe hemorrhage, how and by what mechanisms will the cardiovascular system adapt to reduce the fall in blood pressure?” Our students were good memorizers of anatomic facts about the heart and blood vessels, but this kind of “what-if” question about cardiovascular function was something they were not being prepared to answer. Furthermore, they were not developing the pattern of thought required to understand the body as an interdependent set of self-regulating systems. Developing predictive, what-if thinking became one of our broad course goals.

How do we decide learning priorities?

Textbooks (11, 12) and official syllabi commonly contain a comprehensive content list, and attempting to include all of the topics in a course would necessitate superficial coverage of most. One of the outcomes requested by our nursing colleagues was better problem-solving skills in physiology. We knew that learning the material well enough to develop any useful level of problem-solving ability would require more time on each topic. Because we could not increase the length of the course, we were forced to establish topic and skill priorities to have adequate depth of coverage in critical areas. Again, with the help of our clinical colleagues, we attempted to identify high-, medium-, and low-priority topic areas. This is step 2 of Fig. 1. High-priority areas would receive more time and would include high-level problem solving. Low-level areas would receive either superficial or no coverage and include little or no problem solving. At the HAPS Annual Meeting in 1996 (10), we ran a session in which we asked the 100 or so A&P instructors present to generate their own set of broad A&P course goals (10). Table 1 lists the goals they produced and the number of breakout groups contributing the same goal. The list seemed very close to what our nursing colleagues were suggesting to us.

<table>
<thead>
<tr>
<th>Goal Category</th>
<th>No. of Times Listed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developing thinking, reasoning, data-handling, and interpretation skills</td>
<td>9</td>
</tr>
<tr>
<td>Making realistic applications of learning</td>
<td>5</td>
</tr>
<tr>
<td>Learning and applying the scientific method</td>
<td>4</td>
</tr>
<tr>
<td>Using homeostasis and control as a theme</td>
<td>4</td>
</tr>
<tr>
<td>Integration across disciplines and relating form-function</td>
<td>2</td>
</tr>
<tr>
<td>Developing three-dimensional conceptualizations and relating form-function</td>
<td>2</td>
</tr>
<tr>
<td>Understanding development</td>
<td>1</td>
</tr>
<tr>
<td>Preparing for qualifying exams</td>
<td>1</td>
</tr>
</tbody>
</table>

At the HAPS Annual Meeting in 1996, we ran a session in which we asked the 100 or so A&P instructors present to generate their own set of broad A&P course goals (10). Table 1 lists the goals they produced and the number of breakout groups contributing the same goal. The list seemed very close to what our nursing colleagues were suggesting to us.
taught. When we tallied the responses from the breakout groups, the priorities were, in descending order

1) nervous system,
2) cardiovascular system,
3) fluid balance and renal system and acid-base balance,
4) immune system,
5) pulmonary system,
6) endocrine function, and
7) musculoskeletal system.

This ranking is very similar to what our nursing colleagues gave us but not very similar to the allocation of time and resources in many A&P courses, including the one we began with. That course invested large amounts of time on memorizing facts about the musculoskeletal system, the last-ranking topic.

**Objective Definition**

<table>
<thead>
<tr>
<th>Objective ID</th>
<th>Priority</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Understand the mechanical activity of the heart and the relationship between rate, filling (EDV) and emptying (ESV, EF).</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>Understand the continuity and serial nature of flow between the right heart, lung, left heart and systemic circulation, and the role and of the valves and their action in maintaining this circuit.</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>Understand the relationship between electrical and mechanical events in the heart.</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>Understand the intrinsic and extrinsic regulation of cardiac output.</td>
</tr>
</tbody>
</table>

FIG. 2.

Data entry screen for defining topic learning objectives. Shown are some learning objectives for the cardiac unit. Priority indicates the relative importance of the objective, which determines the amount of time that will be devoted to it.

How do we create course content and pedagogy that is aligned with our priorities? Once we know our broad course goals and topic priorities, we face the challenge of deciding what to cover in each topic and assembling the course components that will enable learning. This is the point at which we start defining learning objectives (step 3 in Fig. 1). We cannot assemble the course without knowing specifically what is to be learned in each topic area. When we begin each new topic area, we start with some basic questions about which learning objectives are indispensable. These determinations are also ideally made in consultation with clinical colleagues and licensing exams. We usually choose 4-10 broad learning objectives for each topic area. Realizing that any curriculum is a highly cross-linked and interrelated set of elements, we developed a relational database to facilitate the design process and help us establish and utilize the web of physiological knowledge we are
teaching. Figure 2 is a data entry screen on which topic (unit) learning objectives for the cardiac unit are defined. These broad learning objectives form the topic outline. Each objective has an identification and a priority that helps us decide how much time it should receive. The priority of the objective is an important feature when we assemble exams.

The term “understand,” which appears in each of the learning objectives shown in Fig. 2, has no clear, operational meaning and could be interpreted in a wide variety of ways. Because of this, we realized that for each of a topic’s learning objectives, we needed a set of specific statements that we called benchmarks (step 4 of Fig. 1) (9). Benchmarks define concretely (operationally) what it means, for instance, in the case of the first objective in Fig. 2, to “understand the mechanical activity of the heart...” One benchmark for that objective is to “predict the cardiac output based on variations in either heart rate, stroke volume, end-diastolic volume, or end-systolic volume.” We chose the benchmark term because its classic definition of being a fixed marker of elevation seemed to denote the clear demarcating character we wanted the statements to have. The ability to predict cardiac output, given the information indicated, is one unambiguous way for students to demonstrate their levels of mastery of this topic. Table 2 shows an example of another cardiac content objective with all of its benchmarks.

On the basis of our initial experience, we developed a series of template model statements as a starting point for writing benchmarks (Table 3) that follow a cognitive/difficulty progression somewhat like Bloom’s taxonomy (2).

Figure 3 is the benchmarks entry screen. The benchmarks shown are related to the objective at the top of the screen. Benchmarks, in addition to having a name and description (derived from the indicated model statements), are assigned one of three difficulty levels. At first we conceived of the model benchmark statements as comprising a difficulty continuum from 1 to 9, with terminology (1) being low level and application and transfer (9) high level. We thought their selective use would allow us to set the difficulty level of a topic area. We then realized that, whereas some model statements always lead to benchmarks at a predictable difficulty level, others may lead to benchmarks at several levels depending on the area and the particular way they are applied. Terminology benchmarks (1) are almost always low level, but comparing and contrasting (4), for instance, could be relatively simple or exceedingly complex, being anywhere from low to high level. Prediction also could be either relatively straightforward and easy or very difficult. Levels are not easy to assign confidently at the outset. In fact, we know of no objective way to determine a priori the difficulty level of a benchmark. We established three possible levels that are loosely based on the concrete, transitional, and formal cognitive levels of Piaget (14). Initially, the benchmark levels are as-

<table>
<thead>
<tr>
<th>Content objective: Understand the relationship between electrical and mechanical events in the heart</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Benchmarks</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
</tbody>
</table>
signed subjectively, and over time they are reviewed and modified if we find them to be more or less difficult than we expected, based on empirical observation of student performance.

Benchmarks also have four possible links to other curricular elements. They may be linked to 1) one or more other units (benchmark 1 in Fig. 3), 2) one or more objectives, 3) one or more common misconceptions (benchmark 2 in Fig. 3), or 4) one or more fundamental concepts (benchmarks 2 and 3 in Fig. 3). Common misconceptions are those we have frequently encountered with students and are specifically addressed by the benchmark (16). An example of a common misconception about the heart is that atrioventricular valve opening and closure is controlled by the papillary muscles. Fundamental concepts are usually processes such as diffusion or gradient-generated flow, which must be understood for the benchmark to be mastered. Linking is an important feature of the benchmarks process, because it makes explicit the interdependence of the many aspects of understanding A&P.

Although the examples of objectives and benchmarks given so far are content related, they also may be oriented toward the development of skills. Broad course goals such as those in Table 1 often point to the systematic development of skills throughout the course. Included are problem-solving ability, facility

| TABLE 3 | Model benchmarks statements |
|-----------------|
| 1) Associate the key _______ terminology with the processes (and structures) of the _______. |
| 2) Identify the (physical) features and locations of the key elements of the _______ stating the way in which each is determined by its most important functions. |
| 3) Trace the movement of _______ through the _______. Identifying the sequence of important structures (and/or events) and explaining how this sequence is necessary for the _______ to function properly. |
| 4) Contrast or compare the _______ of the _______ with the _______ of the _______. |
| 5) Relate the behavior of the _______ to the function of _______. |
| 6) Use the normal range of the _______ to determine physiological state of the _______ system. |
| 7) Demonstrate the ability to _______ (skill name) by _______ (means). |
| 8) Predict (and calculate) how _______ would vary based upon changes or disruption in _______. |
| 9) Apply (Transfer) the concepts of _______ in the _______ (system or concept) to the _______ in the _______ (system or concept). |

Benchmarks also have four possible links to other curricular elements. They may be linked to 1) one or more other units (benchmark 2 in Fig. 3), 2) one or more objectives, 3) one or more common misconceptions (benchmark 2 in Fig. 3), or 4) one or more fundamental concepts (benchmarks 2 and 3 in Fig. 3). Common misconceptions are those we have frequently encountered with students and are specifically addressed by the benchmark (16). An example of a common misconception about the heart is that atrioventricular valve opening and closure is controlled by the papillary muscles. Fundamental concepts are usually processes such as diffusion or gradient-generated flow, which must be understood for the benchmark to be mastered. Linking is an important feature of the benchmarks process, because it makes explicit the interdependence of the many aspects of understanding A&P.

Although the examples of objectives and benchmarks given so far are content related, they also may be oriented toward the development of skills. Broad course goals such as those in Table 1 often point to the systematic development of skills throughout the course. Included are problem-solving ability, facility

| FIG. 3. |
| Data entry screen for defining benchmarks of one of the cardiac learning objectives. Three of the benchmarks for this objective are shown. Each benchmark may be linked to other units and objectives, common misconceptions, and underlying fundamental concepts; the existence of such links is indicated by a check. |
with graphic information, and the ability to interpret data. Specific objectives and benchmarks need to address these goals throughout the various topic areas of the curriculum. One of our skill benchmarks in the cardiovascular unit is “being able to interconvert rates and intervals.”

Once the benchmarks are written we have the basis for assembling the learning activities [labs, class-time material, workshop problems, reading, computer-assisted learning (CAL)]. This is step 5a of Fig. 1. There must be at least one learning activity designed for each benchmark, and no learning activity can be included unless it addresses a specific benchmark. Identifying or developing learning activities are tasks with which we instructors typically spend most of our course development time, and we are concerned that the activities be pedagogically sound. Although the benchmarks process is not about pedagogy per se and is mostly about defining what a course is expected to accomplish, it does provide a way of evaluating pedagogy and sometimes suggests the pedagogical approach. For example, the explicit naming of links to other units, objectives, and underlying concepts strongly encourages the linking of these during the learning process. The student use of concept maps (13) is one pedagogical approach to making sure that the links are noticed and learned.

The link between benchmarks and learning activities, in relational terms, is “many to many”; that is, a particular learning activity, such as a laboratory exercise, may address a number of benchmarks, and a single benchmark may be addressed by more than one learning activity (see Fig. 4). For example, the benchmark of predicting cardiac output from heart rate, stroke volume, end-diastolic volume, or end-systolic volume, as in benchmark 1 of Fig. 4, is addressed in a laboratory exercise, part of a CAL module, and in the textbook reading, providing three distinct learning activities. A cardiac laboratory activity we use in which a quantitative physical model of a single-chambered heart is studied (8), such as the lab in Fig. 4, addresses multiple benchmarks, including benchmarks on valve function, pressure gradients and flow, and the trade-off between heart rate and end-diastolic volume.

One of the pedagogical benefits of this many-to-many relationship between benchmarks and learning activities is that it helps address the issue of student learning styles (4); several different learning activities addressing the same benchmark gives students with different learning styles a range of ways to master it. As learning activities are included in the curriculum, they are explicitly linked to all of the benchmarks they address so that we know where they are relevant. Again, learning activities have to address at least one benchmark to be included in the course, and every benchmark must be addressed by at least one learning activity. If benchmark 1 of Fig. 4 is eliminated, “orphaning” the CAL and text learning activities, those activities would also be dropped because they no longer serve a benchmark. The lab that addresses benchmark 1 would not be dropped, because it still addresses benchmarks 2 and 3, which remain. Stating these connections helps keep unnecessary activities out of the course, such as laboratory exercises that have only marginal relevance. It also ensures that
students have learning activities for everything we expect them to learn.

**How do we communicate to students what their learning objectives should be?** Once we have carefully laid out our curriculum, we need to be sure that students are informed about the learning objectives and benchmarks. This is again a pedagogical consequence of the benchmarks process; almost everything about the course has been made explicit. The objectives and benchmarks are the guide for students about what is important, and, as we will see next, they remove the uncertainty about the nature of examinations. We currently give students the objectives and benchmarks exactly as they exist in the database but have considered rewriting them in a way that is more informative for a learner who is new to the material. The instructor’s and student’s perspectives are often quite different, and the same statement may not convey the same idea to both.

**How do we assess students’ success in achieving their learning goals (benchmarks)?** Figure 1 shows that assessment is derived from the benchmarks (step 5b). In fact, well-written benchmarks are so specific that it is very easy to decide what to assess (5, 7). This unfortunately does not remove the difficulty of generating good assessment items, but it does ensure that they are aligned with what students have done. Alignment is automatic, because learning activities are also derived from the same benchmarks. It is very easy to write an assessment item based on the sixth benchmark for the cardiac objective in Table 2 (“Associate the key features of the ECG with the electrophysiological and contractile processes of the heart”). Essay, matching, multiple-choice, concept mapping, and other types of questions could all be written to assess achievement of this benchmark.

Objectives have priority levels, and benchmarks have difficulty levels. Both of these parameters can be very useful in constructing balanced examinations. High-priority objectives should be allocated more learning time and should be represented in the same proportion on the exam. Because the prioritization is explicit, it can be used to plan the assessment in a way that fairly represents its coverage in the course. It is our intent to construct exams that also demand varying levels of intellectual performance. Some (3) have suggested that the mix of high- to low-level questions could be adjusted as the term progresses and students become more proficient at high-level benchmarks. Knowing benchmark levels allows for the selection of assessment that produces the desired ratio of low- to medium- to high-level questions. We can begin with more low-level and fewer high-level questions at the beginning of the term and then adjust the mix with subsequent exams until the final proportions are reversed.

Closely aligning benchmarks, learning activities, and assessment makes it possible to evaluate the efficacy of learning activities. If we are successful or unsuccessful at achieving learning in a particular area, we can easily identify the learning activities that are responsible and make changes in or replace the unsuccessful ones.

**EVOLUTION OF THE BENCHMARKS A&P CURRICULUM**

Any curriculum should be plastic and evolving both to improve what it does and to maintain relevancy. The explicit relationships between the various elements of the benchmarks curriculum makes it possible to change, add, or delete an objective or benchmark and be aware of how learning activities and assessment will be affected. Ideally, this should help overcome the inevitable inertial barrier to curriculum improvement and make it easier to make the changes when they are needed. Often, more than one instructor is involved in writing curriculum, and the benchmarks approach lends itself to collaborative work because it is easy to assign topic areas to individuals. Agreeing first on broad objectives and priorities makes it easier to make decisions about what learning activities to include. Deriving outcomes as much as possible from information external to the course both maintains relevance and reduces disagreements based on “pet” areas of coverage. Review and revision is easier because the format is always the same and there is a well-developed rationale for everything that is included.

**NATIONAL A&P CURRICULUM DATABASE**

We have just initiated a new stage in the evolution of the HAPS Core Curriculum called Curriculum 2000 (10). To do so, we are adapting the benchmarks model
FIG. 5.
Use of the benchmarks curriculum model for a multicourse curriculum. Dotted lines indicate the flow of information that guides curriculum development. The ultimate source of this information is the destination of students. Desired curriculum outcomes inform the curriculum goals, topic priorities, objectives, and benchmarks. Course 3 is the terminal course in the sequence, and its benchmarks help determine the objectives and benchmarks of course 2, which precedes it. Solid arrows indicate the sequence of developmental steps as in Fig. 1.
so many individuals can participate in the process via the World Wide Web. Defining outcomes will be an important early activity in this process, and we are developing web-based surveys for clinical allied health instructors that will be linked to our curriculum database. Course goals and outcomes can be referenced to these data.

The new HAPS Core Curriculum will include a set of web-based tools to allow members to construct their own individual curricula from its many options and levels, knowing that their selections will have learning objectives comparable to those of others and will be consistent with the consensus of the organization. Over time, the learning objectives will be linked to a resource list of learning activities, and an assessment bank will be developed to evaluate success with the benchmarks. All of these elements will be available via the web to instructors and as parts of an extensive database, which will hopefully continue to expand over the years. The Core Curriculum in this version will become a curriculum resource for instructors, allowing them to design their courses to meet the particular needs of their students and to locate resources they need to meet their own chosen set of learning objectives.

USING THE BENCHMARKS APPROACH FOR MULTICOURSE AND MULTI-INSTITUTION CURRICULA

Most curricula include a number of sequenced and nonsequenced courses. These include core undergraduate science curricula, most allied health curricula, and medical school curricula. The benchmarks approach may be used to design and implement a multicourse curriculum that is referenced to a set of outcomes as shown in Fig. 5. The links of benchmarks to fundamental concepts in the terminal course of the curriculum define some of the learning objectives of the courses that come earlier. Another way to say this is that the output state of one course in the sequence is matched to the input state requirements of the next course in the sequence. The fundamental concepts underlying the benchmarks of the first course in the curricular sequence define part of what students must bring with them when they enter the curriculum. Poor articulation between courses in a sequence is often caused by a lack of specificity in these input and output expectations, and the benchmarks approach provides the needed specificity. In large university systems in which community colleges prepare many students for senior college work, the benchmark approach allows for effective articulation to occur.

CONCLUSION

We have attempted to describe an evolving approach to curriculum development. The process helps us always be aware of what we are attempting to accomplish and gives us the flexibility to make changes when they are needed. We think that the use of a web-accessible benchmarks database model will make it possible for many to collaborate on an ongoing basis as we all work to improve our courses, and this will take curriculum development to a new level. However, this technological sophistication, although useful, is not necessary to successfully employ the benchmarks approach. Also, the specific set of steps in Fig. 1 is not required. Any system that begins with a clear, operational articulation of learning objectives that are based on the knowledge of what students need most to learn will produce the same result: dynamic and relevant courses. The benchmarks demarcate the terrain we intend our students to traverse. The rich set of learning “tools” available to us, both traditional and from contemporary technologies, allows the benchmarks to be reached.

This work was supported in part by National Science Foundation Grants DUE-9354477, DUE-9451852, and DUE-9653034.

Address for reprint requests: D. E. Lemons, Dept. of Biology, J526, City College of New York, CUNY, 138th and Convent Ave., New York, NY 10031.

References


