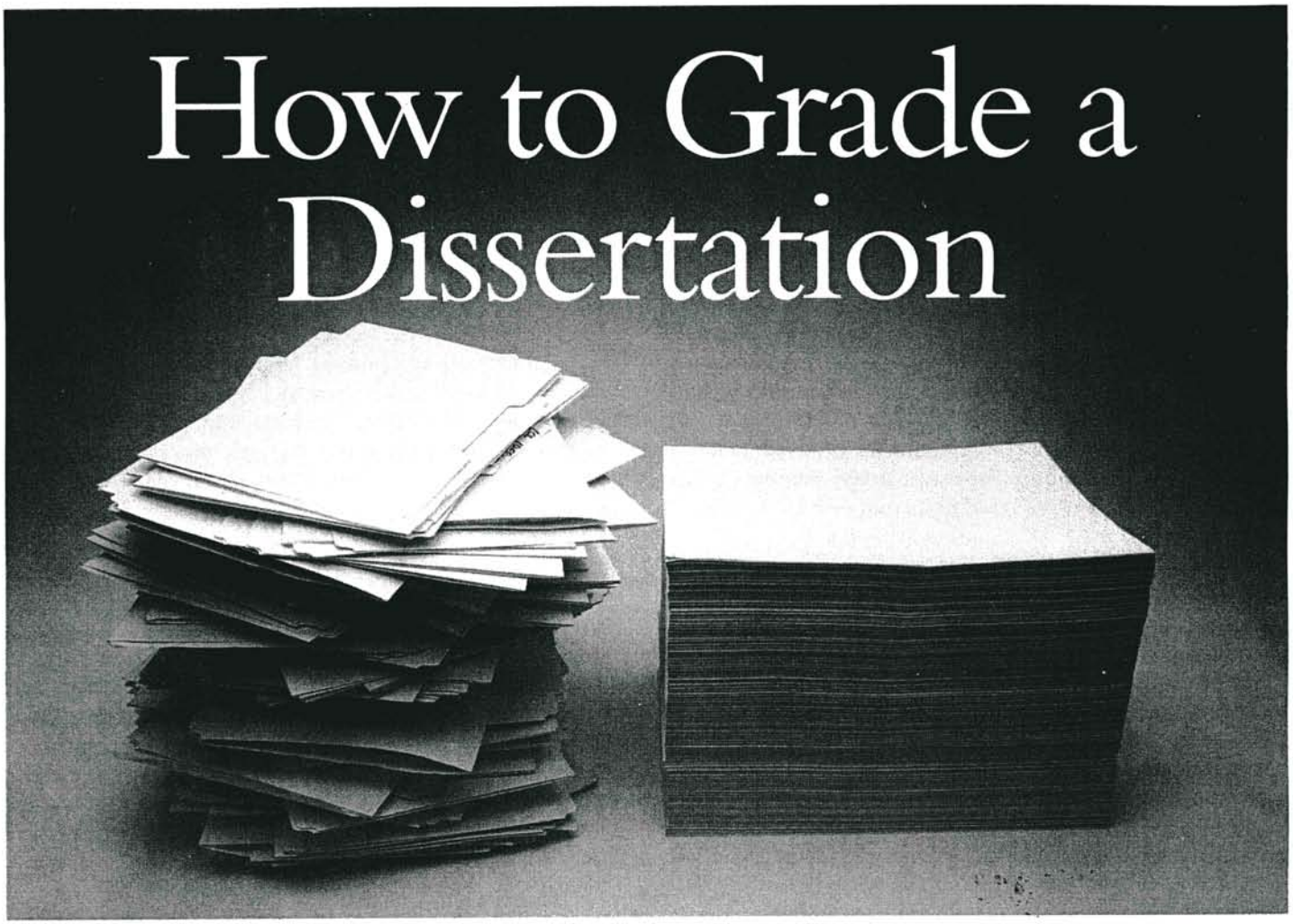
How to Grade a Dissertation



Faculty have implicit standards for evaluating the dissertation. They owe it to their students to make those standards explicit.

By Barbara E. Lovitts

Barbara Lovitts is a senior program officer at the Center for the Advancement of Scholarship on Engineering Education at the National Academy of Engineering. She is author of Leaving the Ivory Tower: The Causes and Consequences of Departure from Doctoral Study. *Her book* Making the Implicit Explicit: Creating Performance Expectations and Assessing the Outcomes of Doctoral Education *is forthcoming from Stylus Publishing.*

The PhD dissertation is the ultimate educational product. It reflects the training of its author and the technical, analytical, and writing skills he or she developed in a doctoral program. Successful completion of the dissertation and the award of the PhD certify that the degree recipient can do independent scholarly work. That much is generally agreed. But who decides what an acceptable dissertation looks like? What are the standards by which faculty evaluate dissertations?

Identifying these criteria—and they do exist, however reluctant faculty are to write them down—could help faculty develop informed measures of learning outcomes. These measures would constitute powerful indicators of the success of research training, provide a method for evaluating PhD programs, and allow more objective comparisons among them. Such standards would also make evaluation of dissertations more valid and reliable across candidates in a department or field.

This article draws on a study that asked faculty to make explicit their implicit criteria for evaluating dissertations. The study aimed to help departments, disciplines, and universities develop objective standards for the outcome of doctoral training—the dissertation—and to use these standards in two ways. At the student level, the goal is to employ them to improve graduate education and training and make it more transparent to students; at the program level, it is to assess educational effectiveness.

The Study

In 2003–04, 276 faculty members in 74 departments across 10 disciplines at 9 research universities participated in focus groups in which they were asked to characterize dissertations and their components (the problem statement, the literature review, theory, methods, analysis, and discussion or conclusion) at four different quality levels—outstanding, very good, acceptable, and unacceptable. They were also asked what it means to make an original and significant contribution in their disciplines and what the purpose of the dissertation is. The study targeted faculty members who had produced high numbers of PhDs in four science disciplines (biology, electrical and computer engineering, physics or physics and astronomy, and mathematics); three social science disciplines (economics, psychology, and sociology); and three humanities disciplines (English, history, and philosophy).

To get a measure of the faculty members' experience and productivity, we asked them to specify how long they had been a professor, how many dissertations they had supervised, and how many dissertation committees they had served on. Together, the 272 focus group members who provided background information had 6,129 years of experience, had advised approximately 3,470 dissertations, and had sat on about 9,890 dissertation committees. The average focus group participant had been a professor for 22 years, advised 13 dissertations, and served on 36 dissertation committees.

The faculty members said that they often make holistic judgments about the quality of a dissertation after they have read it. In other words, they do not have a mental checklist of items against which they assess a dissertation. Still, our results demonstrate that faculty members do make quality judgments about dissertations and that they can (and did) make those judgments explicit. Across focus groups and disciplines, there was a high degree of consistency in the way faculty members characterized the levels of quality we identified (see table 1). However, no dissertation does or can achieve all of the individual benchmarks the group identified. Indeed, taken together, some of the items are self-contradictory.

Outstanding

When we asked faculty members to characterize outstanding dissertations, they often said that such work defied explication, that there was no single feature or set of defining features: "You know it when you see it." Even though outstanding dissertations are rare—they come along once or twice a decade, if that often, the focus group participants said—the faculty members liked talking about this quality level more than any other.

They said outstanding dissertations are characterized by originality, high-quality writing, and compelling consequences. They said such dissertations display a richness of thought and insight and make an important breakthrough. Their clear and persuasive writing provides a glimpse into the mind of the author—you can see how the student is thinking. Moreover, they are a pleasure to read; they are "page turners." When faculty members read outstanding dissertations, they say, "Wow! Why didn't I think of that?" Each individual component of the dissertation is outstanding, and the components are integrated seamlessly throughout the dissertation.

The faculty members described students who produce outstanding dissertations as very creative and intellectually adventurous. They love, and are passionate about, what they are doing, and they display intense curiosity and drive. They leap into new territory and transfer ideas from place to place. Although they often have great advisers with whom they have rich and satisfying intellectual exchanges, outstanding students typically think and work independently, the focus group participants reported. At the same time, an outstanding dissertation can also be a function of luck. The student may simply be in the right place at the right time.

The faculty members said that they often make holistic judgments about the quality of a dissertation after they have read it. They do not have a mental checklist of items against which they assess a dissertation.

Very Good

The focus groups indicated that most of the dissertations they see are "very good," which is the level of quality the faculty members said they expect of most graduate students. Consequently, they had less to say about very good dissertations than about the other quality levels. Very good dissertations are solid and well written, but they are distinguished by being "less"—less original, less significant, less ambitious, less exciting, and less interesting than outstanding dissertations. They display mastery of the field, address the next question or problem in a research program, and are executed competently and confidently. One or more components of the dissertation may not be as strong as the others. The work expands rather than alters the thinking of a field.

The focus group participants said that students who produce very good dissertations show drive and ability. They have good technical skills, but they may not be in control of all the elements of the dissertation. Sometimes, according to the participants, what might have been an outstanding dissertation ends up being "only" very good because the student did not have or take the time to develop his or her ideas. The student may have run out of time or money, had a job waiting, or may simply have wanted to get on with life.

Table 1. The Characteristics of Dissertations

Below are the criteria the focus group members specified for each level of dissertation quality.

Outstanding

- Is original and significant, ambitious, brilliant, clear, clever, coherent, compelling, concise, creative, elegant, engaging, exciting, interesting, insightful, persuasive, sophisticated, surprising, and thoughtful
- Is very well written and organized
- Is synthetic and interdisciplinary
- Connects components in a seamless way
- Exhibits mature, independent thinking
- Has a point of view and a strong, confident, independent, and authoritative voice
- Asks new questions or addresses an important question or problem
- Clearly states the problem and why it is important
- Displays a deep understanding of a massive amount of complicated literature
- Exhibits command and authority over the material
- Argument is focused, logical, rigorous, and sustained
- Is theoretically sophisticated and shows a deep understanding of theory
- Has a brilliant research design
- Uses or develops new tools, methods, approaches, or types of analyses
- Is thoroughly researched
- Has rich data from multiple sources
- Analysis is comprehensive, complete, sophisticated, and convincing
- Results are significant
- Conclusion ties the whole thing together
- Is publishable in top-tier journals
- Is of interest to a larger community and changes the way people think
- Pushes the discipline's boundaries and opens new areas for research

Very Good

- Is solid
- Is well written and organized
- Has some original ideas, insights, and observations, but is less original, significant, ambitious, interesting, and exciting than the outstanding category
- Has a good question or problem that tends to be small and traditional
- Is the next step in a research program (good normal science)
- Shows understanding and mastery of the subject matter
- Has a strong, comprehensive, and coherent argument
- Includes well-executed research
- Demonstrates technical competence
- Uses appropriate (standard) theory, methods, and techniques
- Obtains solid, expected results or answers
- Misses opportunities to completely explore interesting issues and connections
- Makes a modest contribution to the field but does not open it up

Acceptable

- Is workmanlike
- Demonstrates technical competence
- Shows the ability to do research
- Is not very original or significant
- Is not interesting, exciting, or surprising
- Displays little creativity, imagination, or insight
- Writing is pedestrian and plodding
- Has a weak structure and organization
- Is narrow in scope
- Has a question or problem that is not exciting—is often highly derivative or an extension of the adviser's work
- Displays a narrow understanding of the field
- Reviews the literature adequately—knows the literature but is not critical of it or does not discuss what is important
- Can sustain an argument, but the argument is not imaginative, complex, or convincing
- Demonstrates understanding of theory at a simple level, and theory is minimally to competently applied to the problem
- Uses standard methods
- Has an unsophisticated analysis—does not explore all possibilities and misses connections
- Has predictable results that are not exciting
- Makes a small contribution

Unacceptable

- Is poorly written
- Has spelling and grammatical errors
- Has a sloppy presentation
- Contains errors or mistakes
- Plagiarizes or deliberately misreads or misuses sources
- Does not understand basic concepts, processes, or conventions of the discipline
- Lacks careful thought
- Looks at a question or problem that is trivial, weak, unoriginal, or already solved
- Does not understand or misses relevant literature
- Has a weak, inconsistent, self-contradictory, unconvincing, or invalid argument
- Does not handle theory well, or theory is missing or wrong
- Relies on inappropriate or incorrect methods
- Has data that are flawed, wrong, false, fudged, or misinterpreted
- Has wrong, inappropriate, incoherent, or confused analysis
- Includes results that are obvious, already known, unexplained, or misinterpreted
- Has unsupported or exaggerated interpretation
- Does not make a contribution

In experimental disciplines, the dissertation may be very good rather than outstanding because experiments did not work out as planned, or the results were not as crisp and clear as expected.

Acceptable

In discussing the acceptable dissertation, many focus groups distinguished between acceptable dissertations and marginally acceptable ones, although their discussion of the two was often blurred. Participants explained that acceptable dissertations adequately meet the criteria for the award of the PhD, whereas marginally acceptable ones are just barely over the threshold of acceptability—they pass the "gag test."

The participants agreed that acceptable dissertations are somewhat pedestrian and distinguished by being "not very"—not very original, significant, exciting, or interesting. They contain an acceptable amount of solid work that demonstrates that the student can do research. The work is often a highly derivative, small extension of someone else's work. The writing is good enough, but the dissertation is a chore to read. The acceptable dissertation adds little to the field and lacks consequence.

Students who produce acceptable dissertations were said to be functioning close to their capabilities. Although most are bright, they are missing "a certain quality of mind"—they lack

intellectual power and the ability to think like a researcher. They also lack independence and initiative and thus require coaching and hand holding. Their advisers often give them their topics or problems, feed them ideas, and spend much time writing and copy editing their work.

Sometimes, acceptable dissertations are the function of circumstance or bad luck. For example, a student may not have been in residence and, consequently, not received the advice and guidance needed to produce a better-quality dissertation. More often, students rush their dissertations because they have accepted a job or a postdoctoral position or run out of funds. Others have a family to support or simply run out of steam. In experimental disciplines, otherwise good dissertations are considered acceptable when experiment(s) do not work out, and students get null or negative results.

The focus group participants said that advisers and dissertation committees adjust their standards and expectations for students who produce acceptable and marginally acceptable dissertations. The primary consideration is that the student fulfilled the contract. That is, the student worked hard, did what he or she promised to do in the dissertation proposal, and demonstrated competence. The faculty members also take into consideration such "extraneous" factors as their judgments about the person and the type of career the student is planning. Indeed, a few focus groups debated whether the PhD should be awarded solely on the basis of the quality of the product (the dissertation), or whether their feelings about the person and their knowledge of the process the student had been through should play a role in their decision to award the degree. Ultimately, the "hidden criterion" for the award of the PhD is that the student will not embarrass or harm the reputation of the adviser, the committee members, the department, or the university.

Unacceptable

When asked about the unacceptable dissertation, participants balked. They asserted that they rarely, if ever, failed a dissertation, that dissertations of unacceptable quality were seldom allowed to come before a dissertation committee, that students who produced unacceptable dissertations would probably drop out of the program before advancing to a defense, and that it was the adviser's responsibility to prevent unacceptable dissertations from going forward. The faculty members noted that when they did see an unacceptable dissertation, it was usually the adviser's fault; the student had not been given the opportunity to do well. Because so few of the focus group members had ever been on a committee that failed a dissertation, the characterization of the unacceptable dissertation is based primarily on their experience with unacceptable drafts.

Unacceptable dissertations are poorly written and full of errors and mistakes. They are distinguished by "not"—not original, not thoughtful, and not well done. Unacceptable dissertations do not have a good or clearly defined question or problem. They exhibit a poor grasp of the field and either do not use the proper methods or use them inappropriately. Unacceptable dissertations do not yield new or relevant results, and those they yield are often misinterpreted or oversold.

The participants agreed that students produce unacceptable dissertations for different reasons. Most, however, cannot master professional standards and do not have what it takes to be a researcher. Some should not have been admitted into a program in the first place or should have been stopped before advancing to candidacy. Others cannot handle a big project; they do not understand what needs to done. Many cannot or will not take their adviser's advice or criticism into account and, consequently, produce one bad revision after another. Yet others are capable of producing acceptable dissertations but fail to do so because they have taken jobs or otherwise left the university and have not kept up with research in their area. The result is an unoriginal or out-of-date product. Some fail because they push for the defense even though their advisers have told them they were not ready to defend.

In rare instances, when an unacceptable dissertation makes it to defense, dissertation committees seek excuses to pass it. Most defer to the adviser, hold their noses, and vote to pass.

Many advisers simply wait for students whose work is unacceptable to get discouraged and leave instead of proactively terminating them. Some try to disassociate themselves from such students by sending "signals" or by telling them to find another adviser. Others use the defense to get rid of the student. In rare instances, when an unacceptable dissertation makes it to defense, dissertation committees often seek excuses to pass it. They will take into consideration such things as their feelings about the person rather than the objective document. In the end, most defer to the adviser, hold their noses, and vote to pass.

Practical Implications

Faculty members and administrators who want to develop standards by which to judge dissertations must develop their own performance expectations for their local circumstances and translate them into rubrics, or criteria that identify the expected dimensions of a dissertation. The rubrics should include a detailed description of what constitutes different levels of accomplishment for each of those dimensions.

Rubrics can be developed by analyzing existing products (dissertations in this case), or by reflecting on the objectives of a learning task—which is probably easier for a product the size of a dissertation. In their 2004 book, *Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning,* Dannelle

Table 2. Some Dimensions of the Different Components of the Generic Dissertation
The following dimensions emerged from the analysis of the results of the study described in this article.

Component 1: Introduction
The introduction
- Includes a problem statement
- Makes clear the research question to be addressed
- Describes the motivation for the study
- Describes the context in which the question arises
- Summarizes the dissertation's findings
- Discusses the importance of the findings
- Provides a roadmap for readers

Component 2: Literature Review
The review
- Is comprehensive and up to date
- Shows a command of the literature
- Contextualizes the problem
- Includes a discussion of the literature that is selective, synthetic, analytical, and thematic

Component 3: Theory
The theory that is applied or developed
- Is appropriate
- Is logically interpreted
- Is well understood
- Aligns with the question at hand

In addition, the author shows comprehension of the theory's
- Strengths
- Limitations

Component 4: Methods
The methods applied or developed are
- Appropriate
- Described in detail
- In alignment with the question addressed and the theory used

In addition, the author demonstrates
- An understanding of the methods' advantages and disadvantages
- How to use the methods

Component 5: Results or Analysis
The analysis
- Is appropriate
- Aligns with the question and hypotheses raised
- Shows sophistication
- Is iterative

In addition, the amount and quality of data or information is
- Sufficient
- Well presented
- Intelligently interpreted

The author also cogently expresses
- The insights gained from the study
- The study's limitations

Component 6: Discussion or Conclusion
The conclusion
- Summarizes the findings
- Provides perspective on them
- Refers back to the introduction
- Ties everything together
- Discusses the study's strengths and weaknesses
- Discusses implications and applications for the discipline
- Discusses future directions for research

Stevens and Antonia Levi describe a four-stage process for developing rubrics based on reflection. The first stage involves reflecting on performance outcomes, or what faculty want from students. In the second stage, the rubric developer identifies and lists the details of the task and the learning goals. In the third stage, grouping and labeling, he or she organizes the reflections into component skills, which then become the dimensions of the rubric. The fourth stage, application, involves transferring the lists and groupings to the rubric grid and filling in descriptions for each dimension at each quality level. Table 2 illustrates some of the dimensions that emerged from the focus groups' discussion of the components of the dissertation.

Departments or divisions can meet to work out rubrics for dissertations and their components and even for course assignments, seminar papers, or qualifying examinations. The best advice I can offer for working toward shared understandings of expectations about quality is to write everything down while it's being said—on a blackboard or a flipchart, for example. When faculty members see the proposed characteristics against each other, it's easier to debate their merits and move toward a consensus.

Rubrics and Students

Faculty members can use rubrics formatively with students during the research and drafting stages of the dissertation. Ideally, rubrics should be given to and discussed with students at the beginning of the course or program in which they will be used. For the dissertation, optimal times may be during orientation or the proseminar and again in a dissertation preparation course, if the program has one. Many focus group members and doctoral students who participated in a study that supplemented the one described in this article expressed interest in putting performance expectations in graduate student handbooks and on departmental Web sites.

The rubrics help inform students about the expectations of the discipline or profession and its standards for excellence for research and communication. They provide students with benchmarks against which they can judge and revise their work, thus enhancing their ability to assess and correct their own work and that of their peers (and thereby reducing the amount of work advisers and committee members have to do). In short, the more time departments and faculty spend up front explaining and discussing the criteria in rubrics with students, the smoother and easier the process should be for everyone.

In addition to talking to students about the expectations contained in a rubric, advisers and committee members can and should use it to provide feedback to students on dissertation proposals and chapter drafts. As they read a proposal or draft, faculty members can check or circle the level of performance demonstrated on each dimension of the rubric and return it to the student with the draft. Besides reducing the comments an adviser or committee member may need to

write on the draft, the rubric will help students see how their work differs from the target level and will give them information about what they need to do to progress toward the target.

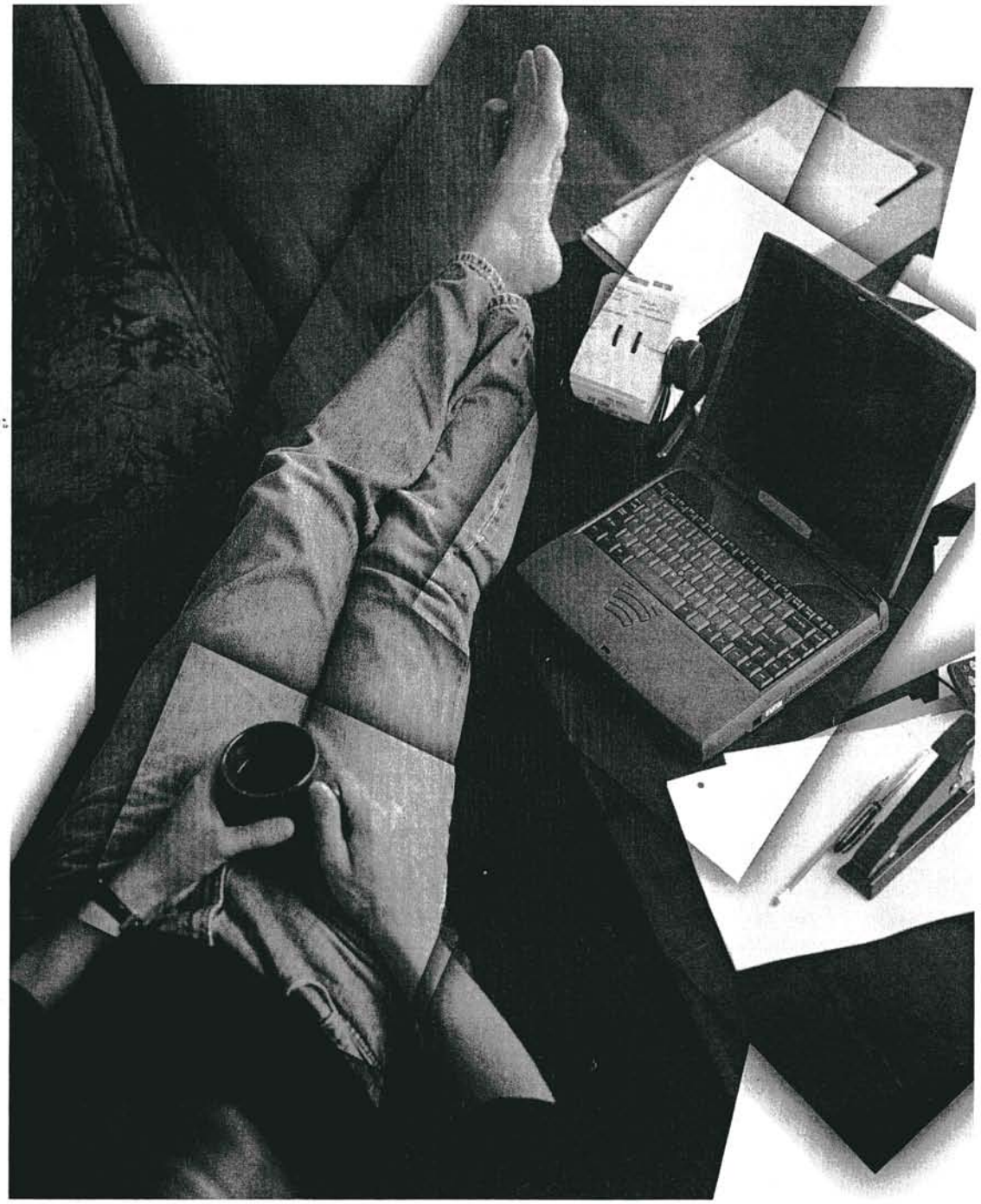
The target should be very good, not outstanding, because very good is the level most faculty members expect of most graduate students. Done is better than perfect. Graduate school is the beginning, not the end, of a career. The goal should be to use rubrics to help doctoral students achieve the highest level possible in light of their needs, capabilities, and professional aspirations, and to get them out of graduate school and into their careers.

Rubrics and Programs

Because rubrics provide a rich record of student achievement, departments, disciplines, and universities can use them to register the aggregate performance outcomes for dissertations for a department, for program or specialty areas, for individual advisers, or for types of dissertations (for example, quantitative versus qualitative, historical, theoretical, or empirical). These aggregate rubrics can offer departments and other stakeholders formal, systematic, evidence-based information about the doctoral program.

As dissertations come up for defense, the adviser, all committee members, or the dean's representative (outside committee member) could be asked to check or circle the demonstrated level of performance on each dimension on a rubric. (Committees should produce a consensus rubric.) The performance on the different dimensions of a dissertation would then be transferred to a master rubric in the form of checks or tally marks.

After the results of a set of dissertations have been recorded on the master rubric, patterns of student performance should become evident. For example, the master rubric might indicate that most students' methodology and analytical techniques fall in the very good range but that fewer students do a good job of aligning theory and methods, or that students who took theory with Professor X do better theoretical work than students who took theory with Professor Y. These patterns should help advisers, departments, and universities identify strengths and weaknesses in many different areas: pedagogy, instructional design, curricular and co-curricular design, institutional programs and services that support student learning, educational resources and tools, educational opportunities, and student advising. This information can then be used to build on the strengths of graduate programs and modify their weaknesses.



Letter to the Editor (submitted)

I was shocked and dismayed when I opened the Nov/Dec 2005 issue of Academe and saw the title that was given to my article, "How to Grade a Dissertation," without my knowledge or consent. The title is antithetical to the intention of my research and represents my deepest fears of how my work might be abused. The goal of my work is NOT to grade dissertations, but to provide faculty and students with a tool they can use throughout the dissertation process to help make faculty's performance expectations for the dissertation more transparent to students and help students achieve to higher levels while they are in the process of researching and writing their dissertations. I wish I had been consulted before such an inappropriate title was given to my article.

Barbara E. Lovitts
National Academy of Engineering