Making the Implicit Explicit

CREATING PERFORMANCE EXPECTATIONS FOR THE DISSERTATION

BARBARA E. LOVITTS

"[This book] offers a sound argument for helping doctoral students achieve high performance levels in the research and writing of their dissertations by providing clear and explicit performance expectations.

The author clearly states that providing . . . explicit expectations should not replace the critical role of the advisor but should enhance the advising relationship between student and faculty member by providing a means for effective formative evaluation. This text is certainly one I wish I had had while writing my own dissertation. In addition to Lovitts' excellent rationales, she gives the reader detailed tables and rubrics that clearly outline the components and characteristics of different quality levels in dissertations.

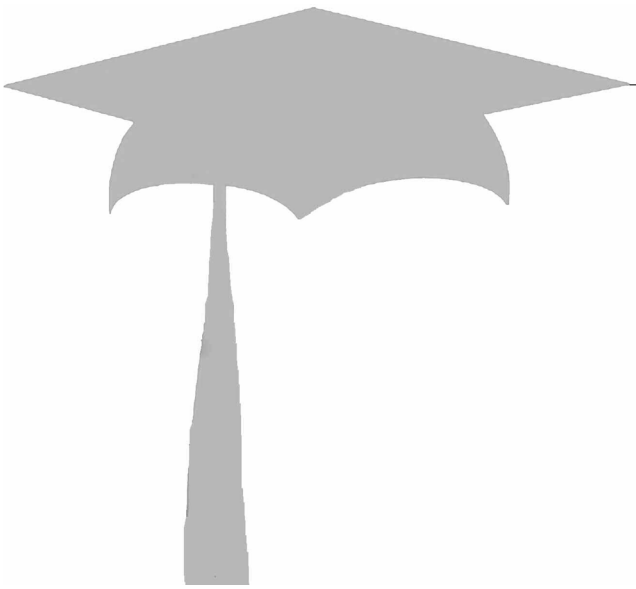
This book is an excellent resource for graduate students beginning the dissertation phase, for faculty who serve on dissertation committees or as dissertation advisors, and for faculty who may teach dissertation process courses. The text is also a valuable resource for academic departments who may want or need to develop dissertation standards from the ground up or to revamp their existing standards and expectations. The strength of Lovitts' book lies in the practical usefulness of the text . . . and its functionality for different academic disciplines.

Students and faculty alike will benefit from this practical and useful resource."

—The Review of Higher Education

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

MAKING THE IMPLICIT EXPLICIT



MAKING THE IMPLICIT EXPLICIT

Creating Performance Expectations for the Dissertation

Barbara E. Lovitts

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

First published 2007 by Stylus Publishing, LLC.
First Edition, 2007

Published 2023 by Routledge
605 Third Avenue, New York, NY 10017
4 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

Library of Congress Cataloging-in-Publication-Data
Lovitts, Barbara E., 1960-
Making the implicit explicit : creating performance expectations for the dissertation / Barbara E. Lovitts.
p. cm.
Includes bibliographical references and index.
ISBN 1-57922-180-7 (cloth : alk. paper) —
ISBN 1-57922-181-5 (pbk. : alk. paper)
1. Dissertations, Academic. I. Title.
LB2369.L685 2007
378.2′42—dc22 2006031897

ISBN 13: 978-1-57922-181-2 (pbk)
ISBN 13: 978-1-57922-180-5 (hbk)
ISBN 13: 978-1-00-344587-6 (ebk)

DOI: 10.4324/9781003445876

CONTENTS

TABLES

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

PREFACE

Too often success in education—kindergarten through doctoral—is a function of "guess my rule." Students who can "psyche out" tests and other assignments or who are socialized into the culture of the assignment and its performance expectations by teachers/professors and/or peers, not only tend to succeed but tend to excel. Even with written feedback, the rest are often left wondering about that special something that they were missing and continue to produce work that could have been better had they known the rule.

This book and the study it is based on is about making the implicit "rules" for the assessment of the final of all final educational products—the dissertation—explicit to doctoral students. The goal of explicating the rules (performance expectations) is *not* to rate or grade dissertations or individual components of dissertations and provide a summary score. Rather, the goal is to make the expectations for the dissertation more transparent to graduate students while they are in the process of researching and writing their dissertations, thereby helping them achieve higher levels. In the language of assessment, performance expectations for the dissertation should be used *formatively*, not summatively, with graduate students to support, not substitute for, the advising process.

This study, like my last one on graduate student attrition (Lovitts, 2001), while addressed to faculty and administrative stakeholders (department chairs, deans of graduate schools, leaders of professional associations, disciplinary associations, national organizations, federal and private agencies, and foundations), is written from a student advocacy perspective. Like my last study, it too fits within a broader context of concern about the quality of doctoral education and the need for reform.

Much of the national discussion about reform has focused on what competencies (knowledge, skills, and experience) graduate students should develop, how they gain such competencies during their graduate education, and whether and how they are appropriately prepared for careers inside and outside academe (Austin, 2002). Since 1990, several national program-level

initiatives have been launched to address these and other concerns. These initiatives include the Graduate Education Initiative (Andrew W. Mellon Foundation), Preparing Future Faculty (Association of American Colleges and Universities and the Council of Graduate Schools with funds from the Pew Charitable Trusts, National Science Foundation, and Atlantic Philanthropies), Re-envisioning the Ph.D. (University of Washington with funds from the Pew Charitable Trusts), The Responsive Ph.D. (Woodrow Wilson National Fellowship Foundation), and the Carnegie Initiative on the Doctorate (Carnegie Foundation for the Advancement of Teaching). Two national student-level studies have also been part of this conversation: Golde and Dore's (2001) *At Cross Purposes: What the Experiences of Today's Doctoral Students Reveal About Doctoral Education*, and the National Association of Graduate-Professional Students' (2001) *The 2000 National Doctoral Program Survey*.

Missing from these initiatives and other reports (Association of American Universities 1998; Council of Graduate Schools, 1997, 2004; Committee on Science, Engineering, and Public Policy, 1995) is a discussion of the role of the dissertation in doctoral education and criteria for its quality, though these were topics of discussion in the early decades of American doctoral education (see chapter 2). Also missing is a discussion of the learning goals of doctoral education, the quality of the educational experience doctoral students are receiving, and how that education prepares students to research and write a high-quality dissertation.

Assessment, which has been a major focus of efforts in the past decade to improve undergraduate education and which is part of some of the abovementioned initiatives, is beginning to receive increased attention at the doctoral level (Borkowski, 2006). Regional accrediting agencies are starting to look more closely at the outcomes of graduate education (Southern Association of Colleges and Schools, 2001). The National Research Council (NRC) (Ostriker & Kuh, 2003) has revised its methods for assessing doctoral programs. And students and taxpayers are demanding greater accountability and demonstrated performance (Borkowski, 2006; Nyquist, 2002).

Although virtually all of the assessments of doctoral education are at the program level, none of the assessments provide formal, systematic, evidenced-based assessment of the quality of doctoral programs (Golde, Jones, Conklin Bueschel, & Walker, 2006). Indeed, the NRC report (Ostriker & Kuh, 2003) states that past measures of educational effectiveness

relied on a question that confounded research reputation and educational quality. The new NRC survey of research-doctoral programs (see www7.nationalacademies.org/resdoc/index.html) seeks, as measures of program quality, information on completion rates, time to degree, placement, profession development opportunities, and so on, but collects no information on the actual educational quality or learning outcomes of the programs.

The Making the Implicit Explicit (MIE) study provides a blueprint for assessing the outcomes of doctoral education at both the individual and program level. As the reader will see, the MIE study found surprising consistency in faculty's characterization of the dissertation and components of the dissertation at the four quality levels they were asked to describe. This consistency suggests that standards can be created for doctoral education writ large and for individual disciplines. Because dissertations reflect the training received, the technical skills, and the analytic and writing abilities developed in a doctoral program (Isaac, Quinlan, & Walker, 1992), such standards would provide stakeholders with a valid and reliable, criterion-referenced measure of student learning outcomes and educational effectiveness.

Having authored another book, I will readily admit that this one was an odd one to write. Each of the ten disciplinary chapters say essentially the same thing. Indeed, the focus-group faculty across universities and disciplines often used the same words and phrases to describe dissertations and components of dissertations at the different quality levels. Therein lies this study's strength. It is only by demonstrating how similar faculty's standards or criteria for evaluating dissertations are in disciplines as dissimilar as engineering and English or physics and philosophy that faculty and administrators across the knowledge spectrum will come to accept the possibility of agreed upon expectations/criteria that approach universal standards. This is not to say that there are not disciplinary differences. There are. English students do not invent devices. Engineers do not analyze and critique texts. These differences should be explored more fully for their disciplinary and transdisciplinary implications. Yet, when it comes to assessing quality, it is safe to conclude that faculty, regardless of university or discipline, do so in very similar terms.

Origins of This Study

The MIE study has its origins in another study funded by the Alfred P. Sloan Foundation I was (and still am) pursuing. That study, The Critical

Transition: From Coursetaker to Independent Researcher (Some Make It, Some Don't) focuses on the disconnect between the criteria by which students are admitted to doctoral programs—being good course takers—and the criteria by which they are awarded the Ph.D.—making an original contribution to knowledge (see Lovitts, 2005). In thinking about the outcomes of doctoral education, I recognized that not all students achieve the goal of originality to the same degree or produce dissertations of equal quality, and that I needed to know more about what it meant to make an original contribution and about how faculty judge dissertations. I combined that need with my basic knowledge of how rubrics are developed for evaluating student performance on high-stakes assessments such as K–12 statewide examinations, the Advanced Placement tests, and the writing component of the Graduate Record Examination, and realized that I could make a contribution to graduate students and graduate education by developing rubrics of performance expectations for the dissertation and its components using a process in which faculty are asked to make explicit their implicit standards or criteria for evaluating the dissertation.

Around the time that I was submitting a proposal for a planning grant for the Critical Transition study to the Sloan Foundation, Chris Golde, senior scholar at the Carnegie Foundation for the Advancement of Teaching, asked me to participate in a panel she was putting together on assessing the outcomes of doctoral education for the 2002 Association for the Study of Higher Education (ASHE) Annual Conference (see http://www.carnegiefoundation.org/CID/ashe/index.htm). I fortuitously had an idea in hand, which I turned into a "conceptual approach" paper (see Lovitts, 2002).

In April 2003, at the Council of Graduate Schools' Ph.D. Completion Project Workshop, my program officer, Ted Greenwood, told me how much he liked that paper and encouraged me to come in with a grant proposal to implement the approach. At that time, I had an affiliation with the University of Maryland and was living solely off the $45,000 planning grant for the Critical Transition and hoping to get a full proposal for that project funded. Fearing a lapse in salary and benefits while the Critical Transition proposal wended its way (unsuccessfully) through review at the National Science Foundation, I decided to pursue the MIE study with a Sloan Foundation Officer's Grant, which goes through review much more quickly than a regular grant. However, Officer's Grants are limited to $45,000 of which about half was available to me for salary and benefits over a six-month period.

Yes, this project was done for $45,000 and a lot of sweat equity, as funds ran out in March 2004 leaving me unemployed and job hunting for several months. I was fortunate to be offered a 60%-time position at the National Academy of Engineering. The initial part-time nature of the position gave me the opportunity to complete the analysis and write this project's results at home two days a week, but without additional compensation.

While working on this and the Critical Transition project, I often scratched my head and wondered why I was underwriting the reform of graduate education with my bank account—and still do. However, there have been a number of priceless intangibles. First and foremost, has been the ability to pursue my own interests in my own time and in my own way. I derive far more pleasure from working on my own projects for my own ends than I do from working on other people's projects for other people's ends. Second, despite the financial hardship, which was considerable, at the end of my life I will not rub my wrists and spirit sore from the shackles of golden handcuffs and regret never having taken a risk. The risk did not pay off as I had hoped, but there is great satisfaction in knowing that work that otherwise would not have been done has and will contribute to a national dialogue and large-scale change. Third, as a single, adoptive mom to a child who was between the ages of two and four during my "quest for independence" (sometimes referred to as "Barbara's folly"), I was able to respond to morning requests of "Read a book" or "Cuddle a few more minutes," because I was in control of my work schedule. I also did not have to kiss and run when I dropped my daughter off at daycare in the morning. I could leave work early and run child-free errands on the way home. And I could take off whenever I wanted to attend an event in my daughter's class. You simply cannot put a price on the benefits we both derived from my being there for my daughter during these formative years—and I'd do it again in a heartbeat.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

ACKNOWLEDGMENTS

This project owes a large debt to many individuals. I would like to thank them in roughly chronological order of their contributions. First comes Ted Greenwood, my program officer at the Alfred P. Sloan Foundation, for his never-ending encouragement and support—both moral and financial. Next comes William Falk, chair of the Department of Sociology at the University of Maryland, for giving me a place to hang my hat and work for over two years. He and the department's office staff—Patty Bernales, Wanda Towles, Gerry Todd, and Joey Irizarry—helped me in many ways large and small. Les Sims, senior scholar in residence at the Council of Graduate Schools, helped identify deans to invite to participate in the study and allowed me to use his name in letters of invitation.

The following faculty: Jane Donawerth, Department of English; James Henretta, Department of History; Christopher Morris, Department of Philosophy; and Jonathan Rosenberg, Department of Mathematics, at University of Maryland all met with me and helped work out the components of dissertations in their disciplines that were used in this study. However, the persistent flaws in the components reside with me, because as experienced practitioners, some of whom are also directors of graduate study, they cautioned me that the components we arrived at were oversimplifications and they did not necessarily fully agree with them. Along these lines, thanks go to Karen Klomparens, dean of the Graduate School at Michigan State University for helping "verify" the components with directors of graduate study at MSU, and for otherwise being an early and committed supporter of the project.

Many thanks go to the deans, coordinators, and facilitators (who were not paid enough) at the participating universities for the time and effort they put into making this study a success. Please see Appendix A for their names and affiliations at the time of the study. Extra special thanks go to Candice L. Miller, director of research and information, who was the coordinator at the University of Colorado site. Candice was an unabashedly enthusiastic supporter of the project with whom I exchanged many, many personal and

professional E-mails during the course of the project and beyond. Candice (Miller, 2006) wrote about the impact of the MIE project at the University of Colorado from an administrator's perspective. Thanks also go to the 276 faculty who generously gave their time to participate in focus groups—I learned a lot from you!

Then there is Jeannie Brown Leonard, my always happy and cheerful graduate research assistant, who has become a good friend. Jeannie did a fabulous job editing and coding transcripts, reducing data—and, yes, listening to and correcting many of the unfortunately very poorly transcribed transcripts. Jeannie also managed to find time in her very busy life as a graduate student, graduate assistant, wife, and mother to do an uncompensated supplementary study that involved getting dissertators' perspectives on the dissertation and its evaluation as well as their reactions to their discipline's MIE performance expectations (see Brown Leonard, 2006).

Special thanks go to Peggy Maki, higher education assessment consultant, for her enthusiastic support of the MIE concept, which she talks about in her travels, and for encouraging me (and Candice, Jeannie, and facilitator Tom Cyr at the University of Colorado) to make a presentation at the American Association of Higher Education's Assessment Conference in Denver in June 2004, for inviting us to contribute chapters to her volume *The Assessment of Doctoral Education: Emerging Criteria and New Models for Improving Outcomes* (2006), for helping me find a publisher for this book, and for advising on revisions.

Susan T. Hill, director of the Doctorate Data Project at the National Science Foundation, discussed with me and provided much of the background data that leads off the disciplinary chapters. Andrew Mary at the National Center for Education Statistics talked with me about and provided the data on graduate student enrollment. A number of people in the disciplinary professional associations supplied background data as well including Michael Gibbons, American Society of Engineering Education; Carla Howrey, American Sociological Association; Jessica Kohout, American Psychological Association; James Maxwell, American Mathematical Society; John Siegfried, American Economics Association; and Doug Steward, Modern Language Association.

During the summer of 2005, John von Knorring, publisher of Stylus Publishing, and I sent copies of the draft manuscript for external review to the MIE deans, coordinators, and facilitators; as well as to assessment and

higher education researchers; and to the executive director or director of education at relevant disciplinary associations. I would like to thank those who provided feedback for their thoughtful comments: Matthew deTemple, Stony Brook; Ronald Ehrenberg, Cornell University; Chris Golde, Carnegie Foundation for the Advancement of Teaching; Jack Hehn, American Institute for Physics; Phillip Katz, formerly of the American Historical Association; Karen Klomparens, Michigan State University; James Maxwell, American Mathematical Society; Candice Miller, University of Colorado; Paul Nelson, American Psychological Association; Rita Nolan, Stony Brook University; Michael Pearson, Mathematical Association of America; Terry Russell, Association for Institutional Research; John Seigfried, American Economics Association; and Andrea Stith, Howard Hughes Medical Institute. The Sloan Foundation also conducted an external review during the summer of 2005. I appreciate the valuable comments from the four Sloan reviewers, one of whom disclosed his identity: Louis Sherman, Purdue University.

The upshot of the external review was that the manuscript needed to be restructured and revised. I am grateful to the Sloan Foundation for providing financial support for the revision process, which included funds for an advisory committee composed of Chris Golde, Karen Klomparens, Terry Russell, and Louis Sherman. The book would not be what it is without their sage advice and moral support.

Not to be forgotten is my editor and publisher, John von Knorring. John offered me a book contract during our initial phone conversation in September 2004, and provided continuous support and advice ever since.

I am grateful to Serena Mann, vice president and general manager of the University of Maryland's television station, whose office was across the hall from mine, for being such a great friend and listening to me vent about project snags and job-hunting woes.

Last but not least, thanks and love go to my mother, Hannah Lovitts, for helping me out financially during my period of unemployment. Similar thanks and love go to my good friends Bob and Joanne Berger for generously watching my daughter on federal holidays, thereby allowing me to get back on track when stomach bugs knocked me off my strict writing schedule, and to my dear friend Chip Cecil for helping out with life in general.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

PART ONE

THE DISSERTATION AND ITS ASSESSMENT

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

I

JUDGING DISSERTATIONS

The Ph.D. dissertation is the final product of years of study and independent research. Successful completion of the dissertation and award of the Ph.D. certifies that the degree recipient has the capabilities and training necessary for independent scholarly work (Council of Graduate Schools, 2004; Isaac et al., 1992). Yet, the standards for evaluating the dissertation are largely unexplicated and mysterious (Holbrook, 2002b; Johnston, 1997). Graduate students are frequently exhorted to make an "original" or "significant contribution" to knowledge. Indeed, this is the most commonly, and often the only, explicitly stated criteria for the award of the Ph.D. (Tinkler & Jackson, 2000; Winter, Griffiths, & Green, 2000). But what "original" and "significant" mean do not appear to have ever been operationalized or objectively defined for graduate students.[1]

Dissertations reflect the training received, the technical skills, and the analytic and writing abilities developed in a doctoral program (Isaac et al., 1992). Thus, knowing the standards faculty use to judge dissertations could lead to the development of informed measures of learning outcomes. These measures would constitute powerful indicators of the success of research training (Holbrook & Bourke, 2002), provide evaluation standards for Ph.D. programs, and allow more objective comparison among them. Such standards would also make dissertation evaluation more valid and reliable across candidates in a department or field. In addition, having explicit standards would enhance faculty supervision and student research performance.

The only "quality" rating of dissertations currently available at most American universities is the recommendation of the dissertation committee.[2] This recommendation is typically to pass or to fail the dissertation, though the dissertation may be passed with a variety of stipulations (e.g., revise, revise and defend again). These stipulations vary by university. However, it is

not clear how well these categories align with the actual quality of the dissertation (e.g., outstanding, very good, acceptable, unacceptable). For instance, one could imagine an outstanding dissertation being passed provided corrections or additions are made because it needed only minor revisions, such as correcting typographical errors, or because a faculty member requested that an interesting finding be explained in greater detail or a novel position on an issue be expounded on further. By contrast, one could imagine an acceptable dissertation being passed without revision because the candidate had accepted a job and was leaving the area or because the dissertation committee found no technical errors and did not feel that further revision would improve the quality of the work. Indeed, evidence suggests that poor-quality dissertations are often passed. Adams and White (1994), in a study that looked at dissertation abstracts, found that a significant number of dissertations that had passed had obvious and sometimes fatal flaws. Johnston (1997) and Mullins and Kiley (2002) found that external examiners of Australian dissertations were reluctant to fail poor-quality dissertations. Thus the question arises: How do faculty judge dissertations? What standards do they apply and how do they discriminate among dissertations of different quality?

Research on Standards and Criteria for Judging Dissertations

Virtually no research exists in the United States on the standards used by faculty to judge dissertations, most likely because dissertation assessment is viewed as a private affair conducted by a committee of experts, and the issue of quality and standards has not attracted the attention of policy makers. However, two studies have some bearing on the issue.

Isaac et al. (1992) surveyed graduate faculty in 10 fields at a large research university to tap their views on the role and function of the doctoral dissertation. Isaac et al. determined that the three major purposes of the dissertation were: (a) a demonstration of the student's research skills, (b) the development (training) of these skills, and (c) making a contribution to the knowledge base of the field. Isaac et al. are careful to point out that "demonstrate" and "train" tended to be mutually exclusive and that the responses came from faculty in fields with different research traditions. "Demonstrate" was most frequently cited by faculty in fields in which solitary research is the norm. It implies that the skills were acquired during graduate training. By

contrast, ''training'' was most frequently cited by faculty in fields where team research and/or laboratory bench research is common. This suggests that the dissertation may be the first experience or part of a continuing training experience.

The faculty were also asked to rate the importance of six characteristics of a doctoral dissertation: independent contribution, originality, significance, substantial time commitment, length of document, and publishable or source of publishable material. Faculty identified independent contribution as the most important characteristic (4.5 on a 5-point scale). ''Originality'' and ''publishability'' tied for second place (4.3). Virtually all disciplinary groups indicated that ''independence'' and originality were essential characteristics of a dissertation. However, Isaac et al. note that it is not clear whether there would be agreement across disciplines on the meaning of these terms.

Adams and White (1994) reviewed dissertation abstracts in public administration, management, planning, criminology, social work, and women's studies to determine the quality of research in public administration relative to cognate fields. They used several indicators to assess quality: existence of a framework of some kind to guide the research, obvious flaws in the research, relevance of the findings to theory or practice, importance of the topic, and an overall indicator of quality. Based on the description of the dissertation presented in the abstract, they concluded that the majority of dissertations either had no framework or might have had a framework but it had to be inferred. The majority of dissertations also either had or probably had obvious flaws, such as a sample that was egregiously too small from which to draw reasonable conclusions, generalizations of findings from a single case study, the use of an obviously inappropriate statistic, the use of a clearly inappropriate research design given the problem or question addressed, blatant errors in logic, or the serious misapplication of some theory to the research problem. Given their assumption that dissertation research should contribute to theory development, Adams and White asked: Was the research relevant to theory? Again they found that the majority of dissertations contributed little, if anything, to the development of knowledge.

Adams and White further ranked the importance of the dissertation topic from ''unimportant'' to ''very important'' and found few that addressed a very important topic. More dissertations addressed a topic of ''average importance'' than an ''important topic.'' Overall, Adams and White's

study calls into question the nature and quality of the standards faculty apply when they review and pass dissertations, as well as the quality of the training doctoral students are receiving if such gross flaws in scholarship can be identified on the basis of the dissertation abstract alone.

The Components of a Dissertation

No studies have been done on faculty's expectations for the different components or tasks of a dissertation (e.g., introduction/problem statement, literature review, theory, methods, results/analysis, discussion/conclusion). However, a very small body of literature exists on the nature, but not quality, of research article introductions (see Swales & Najjar, 1987). And a small but growing body of literature has emerged on students' understanding of and experiences with dissertation literature reviews (Bruce, 1994, 2001) as well as its assessment (Boote & Beile, 2004, 2005; Hart, 1998).

According to Boote and Beile (2004, pp. 6–7), "The dirty little secret known by those of us who sit on dissertation committees is that literature reviews are often (if not usually) inadequate, poorly conceptualized and written, and boring." Using a 12-item scoring rubric, Boote and Beile assessed four randomly selected education dissertations from each of three universities. They found large differences in the quality of literature reviews within and between universities. The best were thorough, critical examinations of the state of the field, while the worst were disjointed summaries of prior research and broad surveys of a haphazard collection of literature. They speculate that these shortcomings stem from insufficient preparation in doctoral programs. Indeed, in a study by Zaprozhetz (cited in Boote & Beile, 2005) faculty ranked the literature review chapter of lowest importance relative to the other chapters of a dissertation, and expected students to complete it alone with little help from faculty.

A View from Abroad

A small body of research exists on how evaluators assess the quality of dissertations in Australia and Britain, countries where dissertations are sent to and read by examiners external to the university (Anderson, 2002; Bourke, 2002; Delamont, Atkinson, & Parry, 2002; Holbrook, 2002a, 2002b; Holbrook & Bourke, 2002; Johnston, 1997; Lovat, 2002; Mullins & Kiley, 2002; Simpkins, 1987; Winter, Griffiths, & Green, 2000). This research has been motivated by policy concerns about the quality and contribution of research

degrees to the national economy, and the effectiveness of the process of preparing researchers (Holbrook, 2002a). The majority of studies have tried to infer examiners' standards for judging dissertations by analyzing the content of their written reports, the language used, and the proportion of space (number of lines of text) devoted to a content category, though interviews and surveys have also been conducted.

Two major themes have emerged from these studies. One is that external examiners' criteria for evaluating dissertations include a variety of technical and indeterminate qualities. The other is that the dissertation must demonstrate originality and make a contribution. Below we explore these themes in greater detail.

Technical and Indeterminate Qualities

Studies that have analyzed external examiners' reports as well as those that have asked examiners about the criteria they use to judge dissertations typically agree that examiners look for a mixture of technical and indeterminate qualities. The technical qualities operate on two levels. On one level are the technical details that reflect the technical knowledge and skill of the candidate. These are things that can be formally taught and objectively assessed, such as command over subject matter and application of appropriate research and analytic methods (Delamont et al., 2002). The other level is more administrative and reflects attention to detail. It has to do with the dissertation's formal structure and presentation, such as appropriate style, clarity of presentation, absence of typographical errors and grammatical mistakes, accurate calculations, and correct referencing and citation (Delamont et al., 2002; Winter et al., 2000).

Indeterminate qualities, by contrast, are qualities that examiners can recognize but not itemize or articulate precisely. Such knowledge is personal and tacit, and as such defies translation into techniques, skills, and formulas (Delamont et al., 2002). These qualities include such things as intellectual grasp, coherence, and critical thinking. They are things that are believed to be "caught" rather than "taught." In other words, possession and display of appropriate indeterminate qualities are part of the candidates' cultural capital and reflect their tacit knowledge of disciplinary and academic culture. This knowledge is acquired through personal experience and interaction with members of the departmental, disciplinary, and university community.

Although the criteria for success or failure of a dissertation cannot and

should not be reduced to a set of written rules, Delamont et al. (2002) note that while the "technicalities must be correct, the real role of the examiner is to judge whether the student has mastered appropriate indeterminate skills and displayed the right indeterminate qualities" (p. 41). Indeed, Mullins and Kiley (2002) found that examiners took a holistic approach toward judging dissertations. They considered the dissertation as a whole and the way that the quality of various parts related to one another. Similarly, Winter et al. (2000) noted that no single dissertation would (or could) be described by all the positive statements provided by the examiners they surveyed. However, they argue that the dissertation should ideally be free of the weaknesses the examiners described, though one or two weaknesses could be compensated for by a wide range of positive features.

Faculty's Standards for Judging Dissertations

The greatest insight into the standards faculty use to judge dissertations come from a variety of studies that either asked examiners about their standards for an acceptable dissertation or attempted to infer standards from their written reports. Winter et al. (2000) surveyed staff at nine institutions in England and received responses from 31 faculty in 21 different disciplines. Respondents were asked to discuss how they differentiated between a passing and failing dissertation. Mullins and Kiley (2002) interviewed 30 experienced examiners in Australia to gain an understanding of what was an acceptable dissertation and the context in which they developed their conceptions. Johnston (1997), by contrast, analyzed 51 examiners' reports on 16 Australian dissertations in a variety of liberal arts and professional fields to learn about examiners' explicit and implicit criteria for judging theses.

The results of these studies are quite consistent. Table 1.1 displays the categories Winter et al. (2000) and Mullins and Kiley (2002) used to organize the commonly identified features of good/passing and poor/failing dissertations. The table presents the terminology used by the authors. However, the order of presentation has been changed in order to align analogous concepts, where possible. Johnston's (1997) findings on good quality dissertations are similar to those of Winter et al. (2000) and Mullins and Kiley (2002).

According to Johnston, the studies that examiners praised:

- were well designed
- were insightful

TABLE 1.1.
Characteristics of Good/Passing and Poor/Failing Dissertations

Winter, Griffiths, and Green, 2000	Mullins and Kiley, 2002
Characteristics of good-quality or passing dissertations	
Coherence	Cohesiveness and clarity
Intellectual grasp	Critical reflection
Originality	Originality of presentation
Presentation	Professionalism (mature comments, accuracy of logic)
Engagement with the literature	Well-structured arguments
Grasp of methodology	
Characteristics of poor-quality or failing dissertations	
Lack of coherence	Lack of coherence
Lack of originality	Work that is not original
Methodological weaknesses	Mixed or confused theoretical and methodological perspectives
Lack of intellectual grasp	Lack of confidence
Poor engagement with the literature	Researching the wrong problem
Lack of generalizability	Not being able to explain at the end of the thesis what had actually been done
Poor presentation	

Note. The data in column 1 are from "The 'Academic' Qualities of Practice: What Are the Criteria for a Practiced-based PhD?" by R. Winter, M. Griffiths, and K. Green, 2000, *Studies in Higher Education*, *25*(1), pp. 32–34. The data in column 2 are from "'It's a PhD, Not a Nobel Prize': How Experienced Examiners Assess Research Theses" by G. Mullins and M. Kiley, 2002, *Studies in Higher Education*, *27*(4), pp. 378–379.

- were well conceptualized
- were carried out competently
- did what they set out to do
- were thoroughly analyzed
- demonstrated a comprehensive knowledge of the literature
- made a significant contribution to the field

Johnston also presents a list of verbatim comments from examiners about their standards for a dissertation. Some of their comments include the following standards:

- Theoretical justification
- Novel concepts, frameworks, or theory
- Innovative or original approach
- Scholarly analysis of foundation concepts
- Suitable for publication

Mullins and Kiley (2002) went one step further than other researchers and asked interviewees what separated a "good" from a "passable" or "standard" dissertation, or, in other words, what made an outstanding dissertation. The one theme that unified interviewees' responses was the use of an artistic metaphor. The metaphor included such terms as "elan," "passion," "excitement," "sparkle," and "elegance." These dissertations were described as having elegant designs and being creative and well sculpted.

Original and Significant Contribution

What counts as an original or a significant contribution to knowledge has not been explicated or codified. Studies that have assessed examiners' reports typically find statements that say that the dissertation is original and has made a significant contribution. A few commentators (Delamont et al., 2002; Isaac et al., 1992) note that what counts as "original" and as a "contribution to knowledge" is discipline specific, but no study has analyzed these concepts along disciplinary lines.

Three studies (Johnston, 1997; Simpkins, 1987; Winter et al., 2000) address the meaning of "original." For Johnston (1997), originality is demonstrated by "the discovery of new facts" or by "the exercise of critical thinking." Simpkins's (1987) study of examiners' reports of education administration dissertations is devoted to an analysis of the way examiners assess critical thinking. Simpkins found "originality" to be a component of critical thinking. The examiners expected dissertations to show originality in two ways: taking an independent line and taking an imaginative approach. Taking an independent line means that doctoral candidates are:

- Making up their own minds when they review ideas and practice
- Identifying trends and issues
- Drawing their own conclusions when assessing the value of a conceptual model, research method, or research material
- Arguing a position (p. 252)

Taking an imaginative approach means that doctoral candidates are:

- Taking a fresh look at questions, issues, and trends
- Being enterprising in interpreting evidence
- Moving beyond the clearly established to the speculative
- Supplementing logic with the intuitive and imaginative, as long as the "intuitively reasonable" is acknowledged as such (p. 252)

Finally, Winter et al. (2000) asked survey respondents who used the term "original" to elaborate on what it "looks like" in the dissertations they examined. Winter et al. provide a long list of their respondents' verbatim comments. They conclude that "originality" refers to such indeterminate qualities as inspiration, responsibility, cognitive excitement, personal synthesis, and to the candidate's "wrestling" or being "adventurous" with the method.

In sum, the existing research on faculty's standards and criteria for judging dissertations is very limited. Few organize respondents' comments into overarching categories, and when they do, the content of the categories remains largely unanalyzed. Further, no study has analyzed examiners' standards or criteria by discipline or component of the dissertation, nor has any study done a fine-grained analysis of the examiners' standards for dissertations of varying quality other than "pass," "fail," and, in the case of Mullins and Kiley (2002), "outstanding." Finally, no study on examiners' standards or criteria has been conducted in the United States where dissertations are assessed by several faculty in the same and related departments who can and do discuss the dissertation and the candidate. Under such conditions, tacit norms for judging dissertations may have emerged.

This book provides the results of a study in which U.S. faculty were asked to make explicit their implicit standards or criteria for evaluating dissertations. The overarching goal of the study, and the process it describes, is for departments, disciplines, and universities to develop objective standards

for the outcomes of doctoral training—the dissertation—and use those standards in a *formative* way to improve graduate education and training, make it more transparent to students, and help them achieve to higher levels. Below is an overview of the study's methods and the structure of the book.

Making the Implicit Explicit

During the academic year 2003–04, 276 faculty in 74 departments across 10 disciplines (*sciences*: biology, electrical and computer engineering, physics/physics and astronomy, mathematics; *social sciences*: economics, psychology, sociology; *humanities*: English, history, philosophy) at 9 Doctoral/Research Universities-Extensive[3] participated in focus groups in which they were asked to characterize dissertations and components (essential tasks) of dissertations tailored to their discipline (see Table 1.2) at four different quality levels—outstanding, very good, acceptable, and unacceptable. They were also asked what it meant to make an original and significant contribution in their discipline, and what the purpose of the dissertation was. (See Appendix B for the focus group protocol.)

Each university hired a facilitator to coordinate and conduct the focus groups, though some universities assigned an administrator to coordinate them. Faculty were selected and invited to participate in focus groups on the basis of their being high Ph.D. productive; that is, they had advised many doctoral students and served on many dissertation committees. It should be noted that Lovitts (2001) found that high producers have different, more positive attitudes and beliefs about graduate students and graduate education and are more academically and socially engaged with graduate students than their low-productive counterparts. Consequently, the comments made by the focus group faculty do not necessarily reflect the views, experiences, or expectations of all faculty.

Most focus groups had three or four participants, a few ran with only two, and one had eight. The faculty were asked to provide background data on the number of years they had been a professor and their experience with dissertations both as an advisor and as a committee member. Many faculty did not know exactly how many dissertations they had advised and virtually none knew how many dissertation committees they had served on. Consequently, they were asked to estimate. When the faculty provided a range (e.g., 25–30), the average of that range rounded to the nearest whole number

TABLE 1.2.
Matrix of Components of the Dissertation Used for Different Disciplines

Discplines	Components					
Biology, physics, engineering, economics, psychology, sociology	Introduction	Literature review	Theory	Methods	Results/data analysis	Discussion and conclusion
Mathematics	Introduction/ problem statement	Discussion of the literature	Statement of results/theorems	Approach to the problem (techniques)	Proof of results	Conclusion/ future directions
English	Introduction (of problem or concept)	Review of sources	Approach to analysis	Justification of chosen texts	Analysis of texts	Conclusion
History	Introduction	Historiographic review	Sources/ methods	Exposition/ analysis	Conclusion	
Philosophy	Introduction/ statement and clarification of the problem	Demonstration of knowledge of the literature	Development/ defense of the thesis(es)	Recognition and response to possible objections	Conclusion	
Outstanding						
Very good						
Acceptable						
Unacceptable/ failing						

was used. Table 1.3 provides background information on the 74 departments and the 272 focus group faculty who provided background data by discipline. Overall, these faculty had 6,129 years of experience, had chaired an estimated 3,470 dissertations, and had sat on a estimated 9,890 dissertation committees. The average focus group participant had been a professor for 22 years, had chaired 13 dissertations, and had served on 36 dissertation committees.

Each focus group session was tape-recorded and the tapes were sent directly to a business services company for transcription. Although each participant's utterances were transcribed separately, no effort was made to link participants individually with their responses. Thus the identity of the speaker is not recoverable from the transcript, neither in most cases is the speaker's gender. (Data on participants' gender and race/ethnicity were not collected.) Consequently, given the national demographics of senior faculty, the overwhelming majority of whom are male, the participants are referred to as "he," unless I listened to that passage and heard a female voice or there is evidence in the transcript that the respondent is female. In most instances where a focus group dialogue is presented in the text in part two, the speakers are referred to as Participant 1, Participant 2, and so on. In a few instances, participants have been given pseudonyms to make the dialogue easier to follow.

Focus group discussion transcripts were edited so that potentially identifying information, such as names, locations, specialty areas, and the like, was altered, taken to a higher level of generality, or deleted. For readability, common but distracting components of speech such as "ah," "um," "you know," "I mean," "I think," and "sort of" were deleted from the quotations that appear in the text unless they were particularly meaningful. Discrepancies in grammar were not corrected. False sentence starts were frequently deleted, as was the word "and" when it linked sentences, as it often does in spoken language. In most instances, ellipses (. . .) are not used to indicate these deletions.

The transcripts were coded by protocol item and component cell. The coded transcripts were entered into N6, a qualitative data analysis software program (QSR, 2002), by discipline for sorting. After each node or item report was printed, each discipline's data were further reduced through a four-stage winnowing process. In stage one, each node report was read, and relevant information was highlighted. The highlighted information was then cut and pasted into an initial summary table by university. In stage two, each

TABLE 1.3.
Background Information on Faculty Who Participated in Focus Groups by Discipline

Department	Number of departments	Number of focus group participants	Total number of years as a professor	Average number of years as a professor	Total number of dissertations advised	Average number of dissertations advised	Total number of dissertation committees participated on	Average number of dissertation committees participated on
Biology	6	21	467	22	243	12	979	47
Mathematics	9	31	758[a]	25[a]	272[a]	9[a]	887[a]	30[a]
Physics	7	25	520	21	290	12	790	32
Engineering	6	24	450	19	350	15	1099	46
Economics	7	33	621	19	417	13	905	27
Psychology	7	28	690	25	454	16	1432	51
Sociology	7	25	567	23	336	13	1016	41
English	7	24	499	21	419	17	829	35
History	9	33	694[b]	22[b]	364[b]	12[b]	1045[b]	34[b]
Philosophy	9	32	863[c]	28[c]	325[c]	10[c]	908[c]	29[c]
Total/Average	74	276	6129[d]	22[d]	3470[d]	13[d]	9890[d]	36[d]

Note. All averages were rounded to the nearest whole number.
[a] Totals and averages are based on the 30 focus group participants who provided data.
[b] Totals and averages are based on the 31 focus group participants who provided data.
[c] Totals and averages are based on the 31 focus group participants who provided data.
[d] Totals and averages are based on the 272 focus group participants who provided data.

initial summary table was read, and the characteristics for that item were extracted and placed in a reduced summary table. In stage three, the characteristics of each item were synthesized across universities into discipline-based summary tables. These summaries, which appear in tables in the individual disciplinary chapters in part two of this book, stay as close to the faculty's actual language as possible. Because of page-length considerations, the characteristics in these tables are not bulleted as they are in Table 1.3 but rather separated by semicolons (;). In the few instances where a matrix cell is empty, it is because that cell either was not discussed in any focus group or the focus groups that discussed a protocol item or a matrix cell did not provide information that was relevant for creating performance expectations. Finally, in stage four, characteristics that appeared in many of the discipline-based summary tables were extracted and synthesized into an overall or "universal" summary for that item. These tables—with bulleted characteristics—appear in chapter 3. The text in chapter 3 and in the disciplinary chapters typically does not repeat what is in the tables, rather the text summarizes faculty's extended and supplementary discussion of the item.

It should be noted that the focus group protocol was initially developed as part of another study and was implemented exclusive of the matrix of components in seven departments (*sciences*: biology, electrical and computer engineering, physics/physics and astronomy; *social sciences*: economics, psychology; *humanities*: English, history) at each of two research universities.[4] Data from these focus groups are not included in the summary tables, though an occasional quote from those focus groups is included in the text when it contributes something important or different from what was contributed by members of this study's focus groups. Those universities cannot be identified, but will be referred to in the text as Public or Private University when one of their faculty members is quoted.

The remainder of part one of this book focuses on the big picture issues. Before presenting the interdisciplinary or synthetic results on faculty's standards and expectations for the dissertation (chapter 3), chapter 2 introduces and explores the concepts of performance expectations, outcomes and performance assessments, and evaluation ethics. It discusses why faculty, administrators, and other relevant stakeholders should care about them. Chapter 4 explores disciplinary differences in doctoral student training and how these approaches contribute to the development of a dissertation that makes an

original contribution to knowledge. Chapter 5 provides advice on how to translate performance expectations into rubrics for formatively assessing a dissertation and for summatively assessing a graduate program. Finally, chapter 6 discusses the study's practical and research implications for different stakeholders (faculty, departments, universities, disciplinary associations, accrediting organizations). It includes caveats about the (mis)use of rubrics when applied to the dissertation.

Part two presents faculty's performance expectations for each of the 10 disciplines in the study. The chapters in this part are organized by domain of knowledge (sciences, social sciences, and humanities) and alphabetically within the domain of knowledge, except for the sciences. Within the science domains, the empirical disciplines are presented before mathematics, a non-empirical discipline. Within the empirical disciplines, the sciences (biology, physics) come before engineering. Each chapter starts with contextual data on graduate education in that discipline.[5]

Readers are encouraged to read their discipline's chapter or the chapter most closely related to their discipline. They are also encouraged to skim some chapters in each of the three domains of knowledge to get a sense of the similarities and differences across domains and disciplines. (There are more similarities than differences.) They are further encouraged to skim other disciplines' chapters because things may be said in other chapters that are true of one's discipline but just happen not to have come up in that discipline's focus groups.

Notes

1. The nature of "original" and "significant" are addressed in a 1997 Council of Graduate Schools (CGS) report. However, this report is rarely cited and does not seem to have had much influence on graduate education. The report mentions that Berelson (1960) addresses these concepts. Although Berelson is widely cited for his work on student attrition and retention, his work on "original" and "significant" has not made it into contemporary discussions about graduate education.

2. In a search of 25 university Web sites for their policies on evaluating dissertations, I found only 3 that required readers to submit a quality rating or a report on the quality of the dissertation.

3. Doctoral/Research Universities-Extensive, according to the Carnegie Classification of Institutes of Higher Education, are institutions that have a wide range of baccalaureate programs and are committed to the doctorate. They award 50 or more doctoral degrees per year across at least 15 disciplines. From http://www.carnegie

foundation.org/Classification/CIHE2000/def/Notes/Definitions.htm (last accessed October 3, 2005, but link is now defunct). The universities that participated in this study are Duke University, Michigan State University, Northwestern University, State University of New York at Stony Brook, Syracuse University, University of Colorado, University of Illinois, University of Kansas, and University of Southern California.

4. Overall, 55 faculty participated in focus groups in this study; 53 provided background data. These faculty had an average of 25 years of experience, advised an average of 15 dissertations, and sat on an average of 36 dissertation committees.

5. The data are ballpark estimates that are based on eyeballing the last few years of available information on graduate and doctoral education. Estimates are provided in order to give the reader a sense of the magnitude of the enterprise rather than precise numbers because the numbers vary from year to year and become dated quickly. The data come from several sources. The data on the number of Ph.D. programs come from Webcaspar (National Science Foundation [NSF], special run, June 2005). The data on the number of subfields in the sciences and social sciences come from the National Opinion Research Center (2004) *2003 Summary Report of the Survey of Earned Doctorates*, as do the data on the annual number of new Ph.D.s. The data on graduate enrollments in the disciplines come from National Center for Education Statistics (NCES) (n.d.), *Digest of Educational Statistics, 2003*, Table 216. The data on the median time to degree come from a special tabulation of the NSF Survey of Earned Doctorates. The data on postdoctoral education come from Hill, Hoffer, and Golladay (2004), *Plans for Postdoctoral Research Appointments Among Recent U.S. Doctorate Recipients*. Some data were supplied or verified by individuals at the disciplinary professional associations. See notes in the disciplinary chapters.

2

ACHIEVING EXCELLENCE

Since the first three Ph.D.s in the United States were awarded in 1861 by Yale University's Sheffield School of Science (Storr, 1953, 1973), the demand for excellence—originality, independence, methodological expertise, substantive contribution, intellectual rigor, textual clarity, and publication[1]—have been forthright. In 1878, Johns Hopkins University President Daniel Coit Gilman, in his annual report to the trustees, wrote that the doctoral candidate should demonstrate "mastery of his subject, his powers in independent thought as well as careful research and his ability to express, in a clear and systematic order, and in appropriate language, the results of his study" (as cited in Goodchild & Miller, 1997, p. 23). Similarly, in a paper titled "The Doctoral Dissertation" by Wilhelm Gardner Hale that was presented at the third annual conference of the Association of American Universities in Chicago in February 1902, Hale asserted that the dissertation must demonstrate the candidate's "power of originating for himself," that the subject matter "be of such scope that it can be treated exhaustively," and that the presentation be organic, clear, and not "*un*literary" (as cited in Nerad & Miller, 1997, p. 76). Remarks about the quality of the dissertation by Princeton Dean Andrew F. West in 1908 (Storr, 1973) indicate that many dissertations deviated from the standards held by early leaders:

> It [the doctor's dissertation] too often exhibits merely the patiently wrought results of a large quantity of mediocre work. . . . It is too often written under the spur of seeking to find something original. This is apt to result in finding something either unimportant or fictitious. . . . Too many theses exhibit merely or mainly power to arrange, classify, and tabulate; too few dissertations show the power to discover, appropriate, and use only

> what is valuable, and to develop a given subject analytically and constructively. . . . In conclusion, I feel that the question of the Doctor's dissertation is a question of quality—the quality of a man's general liberal education—the quality of his subsequent graduate work, and above all his own personal quality as a man of bright, deep, sensible, definite intellectual character. (p. 55)

A perusal of the Council of Graduate Schools' (1997, 2004) policy statements on the purpose and nature of the dissertation and the doctor of philosophy degree indicate that the expectations for the dissertation have not changed much in the intervening years. (See Geiger, 1986, 1993; and Veysey, 1965, for a more detailed history of American doctoral education.)

Despite the early and consistent demand for excellence, doctoral programs have rarely, if ever, been assessed in terms of the quality of the dissertations departments produce. Yet, dissertations provide the most powerful, objective measure of the success of a department's doctoral program, because, as noted in chapter 1, they reflect the training received, the technical skills, and the analytic and writing abilities developed in a doctoral program. Indeed, assessment, when done properly, can help departments achieve excellence by providing insight into a program's strengths and weaknesses.

This chapter focuses on the relationship between assessment and excellence. It describes how the means of assessing quality have changed over time from a focus on inputs and processes to a focus on outcomes, and explores the implications of this change for doctoral education. It then examines the dissertation as an authentic performance, a true test of the goals of doctoral education. Next it discusses the benefits to students, faculty, departments, and institutions of specifying performance expectations for the dissertation. It concludes with a discussion of evaluation ethics.

Assessing Excellence

Because of the stakes involved, various parties have a vested interest in the assessment of doctoral programs. Prospective graduate students want to know where they can get the "best" education. Faculty want to affiliate with "high-quality" programs. Department chairs and deans of graduate schools want their departments and universities to be "highly ranked" so they can attract the "best and brightest" students and faculty (Golde, Jones, et al., 2006).

Yet, excellence in doctoral education has been traditionally defined in

terms of easily quantifiable student, faculty, and department inputs and processes. At the student level, these measures include such inputs as enrolled students' undergraduate Grade Point Averages (GPAs) and Graduate Record Examination (GRE) scores. At the faculty level, these measures include the numbers of hours spent on teaching, number of grants, number of publications, and number of faculty awards. At the department level, these measures include curriculum offerings, credit hours, pass rates on qualifying exams, attrition/retention rates, and time to degree. However, these inputs and process variables are only indirectly related to educational quality and do not guarantee excellence.

In recent years, the public and concerned internal and external stakeholder groups have placed greater emphasis on educational and institutional effectiveness and accountability in higher education (Borkowski, 2006). Concomitantly, attention has shifted from a focus on inputs and processes to a focus on outcomes. Accrediting agencies now stress the importance of articulating observable, measurable learning outcomes for graduating students and of meeting specified outcomes-based standards. Indeed, in 1997 the New England Education Association of Schools and Colleges' Commission on Institutions of Higher Education (Maki, 1998, p. 28) set forth the following expectations in its Student Outcomes Assessment Plan:

> The Commission expects each institution as a means of its dedication to institutional improvement, to monitor its effectiveness in achieving its mission and purposes. Accordingly, the institution collects and analyzes relevant data and uses this information in the institutional planning process as a basis for sustaining quality and self-improvement. Thus, assessment functions as a tool for the encouragement of such improvement as well as a basis for quality assurance.

Outcomes Assessment

Outcomes assessment is a formal, systematic method for collecting evidence about the quality of a program that, in turn, can help faculty and other relevant stakeholders improve the quality of the learning enterprise. It involves specifying the desired learning outcomes (knowledge, skills, attitudes) that are expected to result from the learning experience, assessing the degree to which those outcomes have been achieved, and then making adjustments to the instructional program based on the evidence.