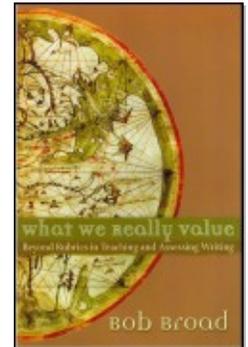# What We Really Value

Bob Broad

Published by Utah State University Press

Broad, Bob.
What We Really Value: Beyond Rubrics in Teaching and Assessing Writing.
Utah State University Press, 2003.
Project MUSE.muse.jhu.edu/book/9267.

→ For additional information about this book

https://muse.jhu.edu/book/9267

2

STUDYING RUBRIC-FREE ASSESSMENT AT CITY UNIVERSITY
Research Context and Methods

For two reasons, readers need to know about the portfolio program at City University and how I studied it. First, the numerous references to sample texts, specific events, and individuals at City University that run through my findings (presented in Chapters 3 and 4) would be difficult to understand and evaluate without a clear picture of both the portfolio program's elements and my grounded theory research methodology. Second, the streamlined version of Dynamic Criteria Mapping I recommend in Chapter 5 for all college and university writing programs is based directly on the full-fledged qualitative inquiry described here. The current chapter therefore describes in some detail both context and methods for my research.

This investigation draws on data I collected in the early 1990s. Those same data have formed the basis for several other studies (Broad 1997, 2000, and 1994b). Though the research questions and the analysis of data in this book differ significantly from those of the earlier studies, the data gathering and most of the theoretical background for my method overlap the earlier work. I have therefore adapted portions of previous discussions of context and methods for use in the current explanation. I encourage readers who seek greater methodological detail to see the earlier studies. [1]

RESEARCH CONTEXT: CITY UNIVERSITY'S
WRITING PORTFOLIO PROGRAM

With the exception of my own name, all names of participants and institutions in this study are fictitious, including the name "City University." To protect participants' anonymity, I have also altered a small number of other nonessential facts about the program I studied. The widespread quoting of words and phrases in this book results from my effort to develop the analysis emically, that is, by drawing it out of the

words and concepts of my research participants. Readers should therefore understand any quoted words not attributed to a specific source (in chapter headings, for example) to be quotations from observations, interviews, or documents at City University.

City University was a large, urban-Midwestern, publicly funded, research university. The College of Arts and Sciences housed the Department of English, which in turn housed the First-Year English (FYE) Program. Nearly all of the roughly four thousand students enrolled each year in the FYE course sequence—English 1, 2, and 3—were in their first year of study at City University. During the fall quarter in which I conducted my study, only three out of fifty English 1 instructors (6 percent) were full-time tenured or tenure-line faculty members; the remaining instructors (94 percent) were adjuncts and teaching assistants (TAs).

The portfolio assessment program I studied was an integral part of the first course in the FYE sequence, English 1. Described in program documents as a course in "writing about personal experience" and in "public discourse," English 1 lasted ten weeks and required students to compose essays on five topics or genres:

> A significant event
> A portrait (often referred to by instructors as the "significant person" assignment)
> Problem/solution essay
> Evaluation (written in a single class period, often taking the form of a review of a cultural event) and
> Profile ("a description of a specific person, place, or activity")

(For details on these assignments, see appendix A, "Assignments for English 1 Essays.") Out of these five compositions, students chose four to include in their final English 1 portfolios. Those course portfolios then served as the basis for a pass/fail evaluation of each student's work at the end of the quarter.

History and Goals of the Portfolio Program

During the ten years prior to implementation of the portfolio program, students at City University were required to pass a timed, impromptu "exit exam" in order to obtain their undergraduate degrees. In the late 1980s, the First-Year English Committee proposed replacing the "exit exam" with a portfolio system, citing a familiar range of shortcomings characteristic of timed, impromptu writing examinations: "the

task is decontextualized; students write in a single mode, at a single sitting; their work of three quarters is subordinated to this single 'testing' circumstance" (*Instructor's Guide,* 95). Two years of piloting the portfolio program yielded extensive adjustments and revisions to address concerns for simplicity, fairness, and time constraints, among others. At the end of the piloting period, portfolio assessment was implemented as a required part of all sections of English 1 (*Instructor's Guide,* 2). In the fall quarter during which I conducted my study (in the early 1990s), City University's portfolio program was beginning its second year of full-scale operation.

The primary goal of the portfolio program was the same as for the timed writing test it replaced: to certify to the rest of the university and to the wider world that students who passed the FYE sequence were "proficient" in reading, writing, and critical thinking. Without that certification, students could not complete their undergraduate degrees at City University. Alongside this certification function, instructors and administrators saw the portfolio program as providing important benefits to the teaching and learning of composition at their university, none of which had been offered by the old writing test:

> Students were judged on the basis of their best work
>
> Portfolio assessment supported best practices in the teaching of writing, including invention, drafting, collaboration, research, and revision
>
> Group meetings and discussions (described below) provided a professional community for instructors within which they felt both safe and stimulated to grow as professionals
>
> Involving outside instructors in evaluative decisions enhanced the fairness and impartiality of the judgments, while also boosting program standards
>
> The portfolio program served a public-relations function, demonstrating the seriousness with which writing was taught and assessed in the First-Year English Program.

Structure of the Portfolio Program

The FYE portfolio program consisted of an intricate system of meetings, discussions, and decisions. Figure 1, "Structure of the Portfolio Program at City University," lays out hierarchical and logistical relationships among the fifty individual instructors and three administrators who staffed the program and among the various groups to which they belonged.

Figure 1
Structure of the Portfolio Program at City University
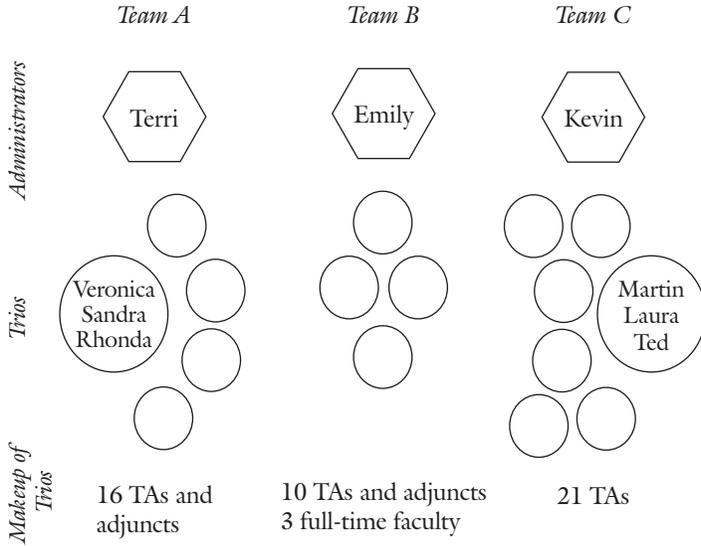
In Figure 1, English 1 faculty are divided vertically into Teams A, B, and C. The administrator leading each team is named in the hexagon at the top of each team. The circles represent instructor-trios, from which I selected two for inclusion in my core group of participants. Names of the instructors in those two core trios, who are quoted extensively in the following chapters, are listed within their respective circles. At the bottom of Figure 1, I provide basic information about the institutional makeup of each team. These figures on numbers of full-time faculty, adjuncts, and TAs do not include the team leaders.

To accommodate different teaching schedules and to limit the size of groups in norming discussions, the twenty-nine experienced English 1 instructors were divided into Teams A and B. Terri, a graduate student in English and assistant to the director of First-Year English, led Team A. Emily, director of First-Year English and a tenured professor, led Team B. The twenty-one teaching assistants (TA's) who made up Team C were all new to teaching English 1 and were all enrolled in the graduate teaching practicum taught by Kevin, associate director of First-Year English and also a tenured professor. Kevin and his practicum students comprised Team C, though Terri led Team C's end-of-term, portfolio

norming session because Kevin was unavailable to do so. While Figure 1 provides a static snapshot of institutional relations among program participants, unfolding events in the portfolio program can best be understood in three dimensions: cycles, modes, and roles.

Two cycles: midterm and end-term Over its twelve-week lifespan, the English 1 portfolio program underwent two nearly identical cycles of evaluative deliberations: a midterm cycle and an end-term cycle. What distinguished the two cycles was that at midterm participants judged English 1 students' individual essays. At the end of the term, portfolios made up of four essays each were the objects of evaluation.

Two modes: norming and trio sessions Each cycle (midterm and end-term) featured group discussions in two modes: norming sessions and trio sessions. In norming mode, each of the three teams discussed the four sample texts found in the *Instructor's Guide,* attempted to articulate their "standards" for evaluation, and voted on—and debated—whether and why each sample text should pass or fail. (I strongly encourage readers to study these sample texts for themselves before reading the following chapters. See appendix B, "Selected Sample Texts from City University.") Norming sessions typically lasted two hours. From norming sessions, instructors went forth and met in instructor trios. Trios collaboratively made the "live" decisions to pass or fail all "borderline" students—those whose instructors had judged the student's work to merit a grade of C or below.

Three roles: teacher, outside evaluator, and administrator Because each student's classroom instructor knew the full pedagogical context within which that student composed his or her portfolio, the teacher played an important, sometimes dominant, role in making evaluative decisions in trios. The teacher's fellow English 1 instructors, who shared the teacher's knowledge of the program and the curriculum, played the role of outside evaluators. Precisely because they lacked the teacher's rich knowledge about a particular student, outside evaluators wielded their own distinctive authority in deciding which students passed English 1. Administrators carefully limited their monitoring of and involvement with instructors' trios, but through the norming process they influenced evaluative decisions, based on their extensive knowledge of the history, theory, and scholarship that informed the portfolio program. (For further discussion of these three complementary forms of evaluative authority, see Broad 1997.)

Chronology of Program Events

Table 1, "Chronology of Key Communal Events in City University's Portfolio Program," allows readers to see precisely how the cycles, modes, and roles described above flowed sequentially.

Where I thought additional detail might prove helpful to understanding my analysis, I have provided it below.

Selection and evaluation of sample texts (preceding spring)   In the spring preceding my fieldwork, the three administrators (Emily, Kevin, and Terri) examined essays and portfolios submitted by instructors as prospective sample texts for evaluation and discussion in future norming sessions. Administrators selected four essays (for midterm norming) and four portfolios (for end-term norming) as sample texts with the stated intention of eliciting certain kinds of judgments and provoking discussion of certain issues in norming sessions the following fall. In interviews, administrators explained that they chose essays to represent a range from "excellent" to "poor" writing quality and to raise such issues as "control of language" or "basic technical competency," storytelling ability, depth of critical thinking, and writers' use of nonacademic dialect.

Preparations for midterm norming (week 3)   In preparation for midterm norming, instructors read the four sample essays ("Anguish," "Pops," "Gramma Sally," and "Belle"), decided whether to vote "pass" or "fail" on each essay, and reflected on how they reached their pass/fail decisions. Meanwhile, the three program administrators met in closed session for about half an hour to review their agreed-upon decisions on each text and to share and rehearse arguments in support of those decisions. Administrators also made an effort to anticipate "trouble" in norming discussions, trying to foresee which texts would likely provoke dissent among the instructors, which instructors were most likely to dissent, and what arguments might arise in support of dissenting evaluations. In these ways administrators prepared to "norm" all English 1 instructors to a clear and cohesive sense of program standards that they hoped would lead to "standardized" (that is, quick and consistent) evaluations in trio sessions.

Midterm norming sessions (week 4)   Norming sessions included four phases: preliminary business, voting on sample essays, discussion of sample essays, and wrapping up. One noteworthy characteristic of preliminary business in midterm norming sessions was an opening statement by the

Table 1

Chronology of Key Communal Events in

City University's Portfolio Program

Chronology			Communal Events
Preceding Spring			Administrators solicited and selected sample texts for use in next year's norming sessions.
Weeks 1 and 2			Preliminary program meetings • Administrators welcomed participating instructors • Administrators introduced themselves, distributed and discussed course materials • Group discussed scheduling, policies, and other topics
Midterm Cycle	Norming Mode	Week 3	Preparations for midterm norming • Instructors read sample essays, made pass/fail decisions, and reflected on their evaluative processes • Administrators met to review pass/fail decisions upon which they had previously agreed when selecting sample texts, to rehearse arguments in favor of agreed-upon decisions, and to anticipate arguments from dissenting evaluators
		Week 4	Midterm norming sessions (Teams A, B, and C) • Preliminary business • Voting on sample texts (pass/fail) • Discussion of voting and sample texts • Wrapping up
		Week 5	Preparations for midterm trio meetings • Instructors evaluated all of their own students' writing ("live" texts) • From their students' texts, instructors selected one "A," one "B," and all those judged at "C" or below and brought those selected texts to their trio meetings ("sampling")
	Trio Mode	Week 6	Midterm trio meetings • One "outside instructor" read all of the teacher's C-or-below texts and indicated whether s/he would pass or fail each text • In case of disagreement between a student's teacher and the first outside instructor, the text was passed to the second outside instructor ("third-reader rule") • Trio members discussed, negotiated, and agreed upon a pass/fail decision for each student text read
Week 7			[None]
End-Term Cycle	Week 8		Preparations for end-term norming (same as for midterm cycle; see Week 3 above)
	Week 9		End-term norming sessions (Teams A, B, and C) (same as for midterm cycle; see Week 4 above)
	Week 10		Preparations for end-term trio meetings (same as for midterm cycle; see Week 5 above)
	Weeks 11 and 12		End-term trio meetings (same as for midterm cycle; see Week 6 above)

leader of the discussion touching upon the purposes, theoretical and historical background, and strategies of "norming." Kevin and Emily delivered such opening statements to their groups; Terri did not.

In each norming session, the administrator-leader wrote the title of each sample text on the chalkboard and then walked the group through the casting and recording of pass/fail votes. On several occasions, instructors resisted administrators' firm directive to "make a [clear] decision" and vote either "pass" or "fail"; some instructors wanted to cast "in-between" votes or to talk about essays before voting on them. Eventually, however, all participants cast a vote one way or the other. Appendix C, "Tabulation of Votes on Sample Texts," displays the votes for the entire program as they were recorded on chalkboards at the end of the various norming sessions and thus indicates levels of consensus and dissent at the end of each norming mode.

The majority of time spent in norming sessions was devoted to discussion of sample texts. At the invitation of the discussion leader, participants volunteered to explain their pass/fail votes. Along with evaluative issues that bore directly upon the decision to pass or fail a particular text, related topics often arose that posed substantial and complex questions or problems for the FYE Program as a whole, such as

> "How do we define 'competency' in English 1?"
> "How important is it for a writer to 'fulfill the assignment'?"

Despite instructors' and administrators' grappling in detail and at length with such pedagogical and evaluative issues in the course of norming discussions, these larger questions were neither addressed nor resolved as decisively as questions about passing and failing particular texts.

End-term cycle (weeks 8 through 12)   As mentioned above, the end-term cycle of norming and trio sessions followed the same pattern as the midterm cycle with one difference: end-term meetings focused upon portfolios of students' writing (four essays including one in-class composition) rather than the single essays discussed and judged at midterm.

This account of my research context begins to answer a central question in qualitative research: "What happened?" To answer the more important ethnographic question "What did it mean?" I employed an array of research methods drawn and adapted from the traditions of qualitative inquiry.

RESEARCH METHODS

Compositionists have begun to explore how teachers and administrators in university writing programs manage the intricacies of shared evaluation (see Allen; Belanoff and Elbow; Reynolds). Still, few studies in the literature of writing assessment systematically analyze communal evaluation as a complex social process, in a specific institutional context, and from multiple points of view within that context. This is the approach I employed to gain insight into what City University's writing instructors really valued in their students' work. The research methods adapted and developed for this study can most usefully be described as an extension of grounded theory (Glaser and Strauss; Strauss and Corbin 1994; Strauss), and, more specifically, as "constructivist grounded theory" (Charmaz). First, I systematically, comprehensively, and recursively analyzed more than seven hundred pages of observational notes, transcripts of group discussions and interviews, and program documents to develop an emic map of City University's terrain of rhetorical values. Working from my best understanding of their experiences, I then brought that conceptual map into dialogue with critiques of traditional writing assessment—and especially of rubrics and scoring guides—current in the literature of evaluation. Extending grounded theory in this way, I found participants' complex criteria for evaluation cast in a new light, suggesting new possibilities for improving communal writing assessment, professional development, and student learning in the classroom. Dynamic Criteria Mapping seizes on those new possibilities.

Background of Research Questions

The good news about the research question driving this study (What did they really value?) is that it did not occur to me until long after my data collection and earlier analyses of those data were complete. From the standpoint of qualitative methods, this late blooming is a good thing because it means this research question could not have inappropriately guided decisions I made in collecting data.

The delayed appearance of questions regarding participants' criteria for evaluating rhetorical performances is also bad news. It means that I did not have the opportunity to interview participants on the most pressing issues in this study, for my interviews were focused on questions appropriate to the analytical frameworks of my earlier research (see Broad 1997; 2000; 1994a; 1994b).

Even though my interview questions focused elsewhere, my interest in new ways to map criteria for evaluation had taken root even before I conducted my City University study. While participating in a variety of communal writing assessment programs during the 1980s and 1990s, I noticed how all judgments were referred back to the scoring guide or rubric. If a criterion did not appear on the rubric, it was discounted as a basis for judging a text; if a criterion did appear on the rubric, it was privileged even if it bore little relevance to the text under scrutiny. This is, of course, exactly how rubrics are intended to work, since their chief purpose is to constrain the range of criteria and thereby boost inter-rater agreement. In these contexts I started to reflect on how rubrics are composed and disseminated as well as how they shape shared evaluation; I began to speculate about what a more detailed, responsive, context-sensitive, accurate account of a writing program's rhetorical universe might look like.

Sometime after my earlier City University inquiries were completed, it dawned on me that this writing program, because it used no scoring guide or rubric, was the ideal setting in which to study criteria for writing assessment in a new way. City University's unconstrained (or, in Diederich's terms, "uncontrolled") evaluation invited instructors to tell unfettered truths about what they valued in the texts before them and compelled others to listen to those truths without dismissing them. In this different atmosphere for communal assessment, participants not only articulated a marvelously rich array of rhetorical values, they also consistently tied their values in shared evaluation settings to their values in first-year composition classrooms. Thus was the analytical approach of this study born, seeking to reveal what instructors really valued in their students' texts.

Establishing and Focusing the Inquiry

In the summer preceding my fieldwork I met with Emily and Kevin (director and associate director, respectively, of the First-Year English Program at City University), and they granted me access to the program for the purposes of my fieldwork. During the program's two preliminary meetings for instructors late in September of that year, I requested the informed, written consent of all participants in the program. All three administrators and forty-eight of fifty instructors granted me their consent to participate. Reservations expressed by two instructors regarding how and whether they wished to participate guided my choices in subsequently narrowing the field of potential participants to a core group.

Defining the core group of participants

As explained above in "Research Context," participants in the FYE portfolio program included fifty instructors of English 1, divided into Teams A, B, and C. During midterm norming (week 4), three factors led me to exclude Team B from the core group of participants:

1. Two members of Team B declined to participate in the research and withheld consent to be recorded or interviewed.
2. Team B had a relatively high proportion of adjunct instructors, several of whom informed me that they simply did not have time to be interviewed.
3. Because of their direct contact with Emily, director of First-Year English, as the leader of their norming sessions, some members of Team B felt certain political risks with particular acuteness.

Once I had limited my main research focus to Teams A and C, I narrowed further my core group of participants by eliminating trios that were incomplete during midterm norming and a trio one of whose members said she didn't have time to be interviewed. Employing what Strauss and Corbin (1998, 201) term "theoretical sampling," I also actively sought out differences in temperament, conduct, and institutional position among the remaining trios. By the end of midterm trio meetings (week 7), this sampling process yielded a core group of four trios—two from Team A and two from Team C—in addition to the three program administrators. Later, during data analysis, I narrowed my core group further, to Trio A-1, Trio C-6, and the three administrators. Quotations presented in my two chapters of findings (chapters 3 and 4) come from these nine individuals in my core group as well as from other members of the norming teams to which my core group belonged.

DATA COLLECTION

I gathered data on the FYE Portfolio Program by observing group meetings, interviewing individual participants, and studying program documents of various kinds. In the actual process of my inquiry, these activities were often juxtaposed, interwoven, and overlapping. To clarify and explicate my methods, this section separates out and details each of these three data-gathering activities.

Observations

I observed and recorded three kinds of group events: norming sessions (six) and trio meetings (fourteen), at both midterm and end-term, and an administrative meeting in preparation for end-term norming. Each

time I observed an event, I made two kinds of records: observational notes and audio or video recordings (or both). I recorded observational notes on every behavior that I thought might turn out to be significant to my analysis. Verbatim transcripts from audio and video recordings expanded my analysis beyond what observations and field notes alone would have made possible and enabled me to study important themes from multiple points of view.

Interviews

Through twenty-seven formal, semi-structured interviews and several informal, impromptu interviews with individual participants, I discovered, developed, and tested interpretations of group events and themes of analysis. I began each interview with an open-ended invitation to the interviewee to reflect upon the events of the portfolio program, such as "As you look back on last quarter's portfolio program, what stands out in your mind?" By starting with such "grand tour" (Agar 1980) questions and following up on them, I was able to build each interview around topics and concerns introduced by participants rather than by me, the researcher. Later during the same interviews, and in second and third interviews where possible, I often asked prepared questions on topics that had arisen from my observations and analyses up to that point. I also posed questions about particular passages in transcripts, inviting the interviewee to look over the transcript before answering the question. This process allowed me to get participants' commentaries on and interpretations of the details of previous events to inform and guide my own analyses. Appendix D provides a sample schedule of interview questions. However, because the research questions for this study only emerged later, I did not pose interview questions focused on criteria for evaluation. Fortunately, some useful material emerged from interviews anyway. One prime opportunity for future research would be for a researcher to study a program's values as I did, but also to formulate interview questions asking participants what they meant when they invoked a particular criterion and why those criteria matter to them. Such an interview-focused study would, I expect, provide extremely useful insights beyond what I was able to learn.

Documents

Because documents powerfully shaped events and decisions in the program, they formed another important source of data. Documents collected and analyzed included:

The mission statement for the First-Year English Program,

The program's *Instructor's Guide,* which included the sample papers and port-
folios discussed in norming sessions and which guided instructors
through the portfolio system,

Drafts and Breezes, a selection of essays from previous years' first-year English
students,

The St. Martin's Guide to Writing (Axelrod and Cooper) and *A Writer's Reference*
(Hacker), the two official required texts for English 1,

Several current students' portfolios discussed during trio meetings

Course readers (collections of articles from the literature in composition and
rhetoric) for the two graduate courses attended by the first-year graduate
TAs participating in my study, and

Various memos, forms, and information sheets distributed by administrators
to instructors.

In helpful contrast to the hurried, day-to-day decisions and actions
squeezed into the ten-week academic quarter during which I made my
observations, these written materials revealed some participants' more
reflective, more historically and theoretically grounded perspectives on
the events of the program. At various times during the quarter, partici-
pants also referred explicitly to program documents in advocating par-
ticular judgments of specific students' texts. For both these reasons,
program documents provided important points of triangulation with my
other data.

DATA SELECTION AND ANALYSIS

My analyses of data passed through three overlapping but distin-
guishable phases. The first phase, *concurrent analysis,* took place during
the twelve weeks of on-site fieldwork and involved reading the incoming
data to guide me in gathering new data. Phase two covered the first sev-
eral months of work on the current study and involved *comprehensive
analysis* of all transcripts and documents to identify, define, and develop
all criteria for evaluation mentioned by my participants. Strauss and
Corbin (1998) call this stage of data analysis "open coding," and Guba
terms it the "discovery mode." In the final phase, *close analysis and verifi-
cation* (named after what Guba calls "verification mode" and closely
related to Strauss and Corbin's "selective coding"), I tested and refined
my Dynamic Criteria Map for City University and other findings (see
chapters 3 and 4).

Concurrent Analysis

The events of the twelve-week portfolio program rushed rapidly past all of us—participants and researcher—and then were finished. To keep myself afloat on that stream of events, I conducted concurrent analysis (formulation of interview questions based on recent events in the program) on the basis of field notes, notes from (not transcripts of) tape recordings, and memories of events available to interviewees and to me, the researcher. The first twelve weeks of the study were, therefore, heavier on data collection than on data analysis.

Interviews integrated questions about and reflections on recent events from both participants and the researcher. One example was Kevin's impromptu suggestion that I interview him immediately following Team C's midterm norming session. His conflict with a first-year TA named Ted over the evaluation of the sample essay "Anguish" made Kevin want to talk about that event immediately after it occurred, so we did. (In a later interview I asked Ted his view of this same conflict.) Working from statements recorded verbatim in field notes, I also asked interviewees about their own or others' particular statements or actions.

As I explained above, my interest in studying the criteria by which instructors and administrators at City University judged students' writing emerged long after my earlier analyses of the data I collected. After deciding to conduct a full-scale qualitative inquiry into these evaluative dynamics, I pursued a fresh course of the second and third phases of data analysis.

Comprehensive Analysis

As I began this phase of treating the data, I searched my entire data pool (approximately seven hundred single-spaced pages of transcripts and documents) for every comment or passage in a transcript that would help answer the question "Why did that participant pass or fail this text?" (because she found it "clichéd," for example) Using QSR Nvivo software for computer-assisted qualitative data analysis, I coded every passage mentioning criteria for evaluation, defined as "any factor that an instructor said shaped or influenced the pass/fail decision on a student's text." My initial list of criteria included 124 distinct factors drawn from approximately 18,500 lines of transcripts specifically devoted to discussing one or more criteria.

Charmaz provides a useful summary of the "constant comparative method," which lies at the heart of data analysis as practiced in constructivist grounded theory.

> The constant comparative method of grounded theory means (a) comparing different people (such as their views, situations, actions, accounts, and experiences), (b) comparing data from the same individuals with themselves at different points in time, (c) comparing incident with incident, (d) comparing data with category, and (e) comparing a category with other categories. (515)

Readers of the current study will see evidence throughout my findings of all these comparative activities.

With some vexation I note that comprehensive analysis—which involved literally hundreds of hours of searching, sifting, sorting, and cross-checking—can be aptly summarized in a couple of brief paragraphs. I console myself, however, with calling readers' attention to how intensively qualitative researchers process data in order to generate conceptual insights.

Close Analysis and Verification

The coding and verification process during this final phase of data analysis refined my interpretations in useful ways. I developed my three levels of grouping criteria (evaluative criteria, criteria categories, and constellations) through repeated and close examinations of the data, and I modified categories extensively in response to what I found there. For example, I found that my list of textual criteria fell into two distinct areas: textual qualities and textual features. Textual criteria also dropped from eighty-two to forty-seven in number. This refinement and condensation of criteria took place when I recognized that some criteria from the earlier list needed to be merged to reflect participants' evaluative frameworks as accurately as possible. For example, "Dialogue" was merged with "Detail and Description"; I merged "Tone" into "Style, Concision." References to "comma splices," which I had originally categorized under "Punctuation," turned out upon reflection and research to be actually a part of the criterion "Grammar." Through this process of refinement I met the qualitative researcher's responsibility to encourage the data to "resist" the researcher's presuppositions and expectations (see Agar; Flower; Guba; Miles and Huberman; Patton) by actively pursuing alternative interpretations of events and attempting to account for all data.

Sifting carefully through large coding reports on each criterion, I first developed my understanding of those individual criteria and then discovered many interrelationships among criteria. Repeated passes through the coding reports yielded more and more finely distinguished, richly illustrated, and complexly interrelated understandings of criteria that eventually served as the structure and content for the Dynamic Criteria Map discussed below.

This account of the theoretical bases and qualitative methods I drew upon in collecting and analyzing my data should help my readers to judge the validity, or persuasiveness, of the analyses and conclusions presented below and to succeed when they try Dynamic Criteria Mapping in their own writing programs. What follows is my exploration and mapping of the criteria according to which participants made their pass/fail judgments of students' texts at City University.