Veronica Boix Mansilla
Elizabeth Dawes Duraisingh
Christopher R. Wolfe
Carolyn Haynes

Targeted Assessment Rubric: An Empirically Grounded Rubric for Interdisciplinary Writing

At the dawn of the twenty-first century, the American academy is marked by renewed interest in interdisciplinary research and education. Multiple drivers propel the growth. Socio-environmental challenges such as mitigating climate change or eliminating poverty demand interdisciplinary solutions. Technologies have ignited interdisciplinary innovations, from unprecedented information sharing to systemic accounts of gene regulation. Recent analyses of the future of industry and labor call for individuals who can understand, employ, and integrate knowledge, methods, and approaches, as well as collaborate across industry sectors, cultures, and disciplinary teams (Levy & Murnane, 2004; National Academies, 2005).

Recognizing this state of affairs, American colleges and universities have increased their interdisciplinary course offerings. In the 2006 *US News & World Report* college and university rankings, 61.71% of liberal arts institutions reported offering interdisciplinary studies majors. In a recent Social Science Research Council survey of 109 American Baccalaureate College-Liberal Arts institutions, 99.07% report either being *very* or *somewhat* oriented to interdisciplinary instruction. In this sample, 65.42% expect to increase their offerings over the next five years

We are grateful to the students at Miami University who contributed their work for analysis. We thank the Atlantic Philanthropies for their generous support of Boix Mansilla and Dawes Duraisingh's work.

Veronica Boix Mansilla and Elizabeth Dawes Duraisingh teach at the Harvard Graduate School of Education. Christopher R. Wolfe and Carolyn Haynes teach at Miami University.

The Journal of Higher Education, Vol. 80, No. 3 (May/June 2009)

(Rhoten, Boix Mansilla, Chun, & Klein, 2006). The American Association of Colleges and Universities (AACU) has called for a renewal of liberal education competencies reminiscent of interdisciplinary learning, such as "integrating knowledge of various types and understanding complex systems; resolving difficult issues creatively by employing multiple sources and tools; [and] working well in teams, including those of diverse composition" (National Leadership Council for Liberal Education and America's Promise, 2007). Among federal funding agencies, the National Institutes of Health (NIH) Roadmap Initiative promotes "interdisciplinary research teams of the future" (NIH, 2006), and the National Science Foundation (NSF) advocates "investigations that cross disciplinary boundaries and require a systems approach to address complex problems" (NSF, 2006, p.6).

Yet the ongoing growth of interdisciplinary programs and courses comes with deep uncertainty about how to structure interdisciplinary learning experiences and measure their success. Overwhelmingly, interdisciplinary programs rely on student grades and opinion surveys (Rhoten et al., 2006). An analysis of four well-regarded interdisciplinary programs (Boix Mansilla, 2005; Boix Mansilla & Dawes Duraisingh, 2007) showed that innovative methods to assess learning outcomes (e.g., real-life problems, portfolios) are informed by generic criteria (e.g., logic of argument or effort and commitment). Such criteria sidestep the question of what, if any, are the defining qualities that characterize interdisciplinary achievement (ibid). In an era of increased accountability, reliable approaches for assessing interdisciplinary learning are necessary to ensure not only the effectiveness of interdisciplinary courses and programs but also their survival (Astin, 1993; Banta, 2002; National Academies, 2005).

A growing body of research on assessment has yielded a plethora of principles and artifacts to monitor and support student learning. Performance-based rubrics, protocols, and portfolios suggest *how* to make the learning contract between faculty and students clear and student learning visible. Yet with few exceptions (e.g. Wolfe & Haynes, 2003; Boix Mansilla & Dawes Duraisingh, 2007), the question of *what* exactly to assess when student work is interdisciplinary remains unanswered. What constitutes quality interdisciplinary student work and how can faculty validly and reliably distinguish between higher and lower achievements? How can administrators discern whether students are developing competencies of interdisciplinary inquiry and communication?

Here, we introduce the *Targeted Assessment Rubric for Interdisciplinary Writing* (Appendix A), an empirically-tested instrument designed to assess interdisciplinary writing at the collegiate level. Interdisciplinary

nary writing presents unique challenges to students, calling upon them to mediate the rhetorical, theoretical, and methodological differences inherent in multiple disciplinary discourses. The rubric proposes four distinct dimensions to be examined: a paper's *purposefulness*, *disciplinary grounding*, *integration,* and *critical awareness*. For each criterion, four qualitatively distinct levels of student achievement are described: *naïve, novice*, *apprentice,* and *master*. The rubric builds on a clear definition of interdisciplinary work, a related assessment framework, and recent scholarship on interdisciplinary writing (Boix Mansilla, 2005; Boix Mansilla & Dawes Duraisingh, 2007; Boix Mansilla, Miller, & Gardner, 2000; Wolfe & Haynes, 2003).

Systematic analysis of student work enabled us to test the rubric's reliability at capturing differences in performance at three stages of collegiate training (freshmen, sophomores, and seniors). These students were enrolled in the Interdisciplinary Studies major at Miami University, following a sequence of interdisciplinary courses in preparation for an individualized interdisciplinary specialization. The rubric is designed as a dynamic tool that researchers and faculty can adapt to examine student work at various disciplinary crossroads, from first-year essays to senior coursework and theses. Aside from grading, the rubric can be used to identify qualities of students' interdisciplinary understanding and support their further development.

Below, we review the assessment literature and the rubric's conceptual foundations. We introduce the rubric through an example of student work and describe the methods by which we developed and tested it. We conclude with concrete recommendations for practice.

Background

Assessing Learning Outcomes

A key marker of institutional effectiveness—despite being difficult to measure—is the quality of individual student learning (Chun, 2002). The drive to advance valid measures of such learning has yielded a range of approaches toward assessment (Chun, 2002; Ewell, 1991; Hutchings, 1990; Schneider, 2002), and assessment experts have advocated the use of rubrics in pre-collegiate and higher education contexts. First, grading is seen to be fairer and more consistent when assessment criteria are made explicit and instructors describe different levels of performance. Second, self-assessment is valued as a means to help students reflect on their work; rubrics allow students to judge the current quality of their work and the ways in which they could develop it further (Brough & Pool, 2005; Huber & Hutchings, 2004; Walvoord & Anderson, 1998).

Some critics charge that rubrics promote shallow learning and are incongruous with student-centered teaching practices because they promote conformity and standardization (Kohn, 2006; Wilson, 2006). Wilson believes rubrics "violate" the complexity of a piece of written work by dividing it into separate, quantifiable parts which do not capture a piece's overall impact or quality. However, as Goodrich-Andrade (2006) points out, some of the perceived shortcomings of rubrics stem from a narrow interpretation of rubrics as tools for grading rather than supports for understanding. She and others (e.g., Huba & Freed, 2000) caution that in a well designed rubric, scoring highly on all of a rubric's criteria is incompatible with not doing the task well. In other words, the power of a rubric rests on the degree to which it captures meaningful dimensions of the work without which a quality product could not be achieved. As suggested earlier, while the authentic assessment movement has broadened the *ways* in which students are assessed, determining *what* to assess has proven more difficult (Boix Mansilla & Dawes Duraisingh, 2007; Cizek, 2000). This is particularly true of interdisciplinary work where a range of epistemological challenges are involved and faculty from different disciplinary backgrounds may have conflicting priorities for students' learning (Klein, 1996; Schilling, 2001). Essential to assessing the *what* of interdisciplinary learning is a clear definition of interdisciplinary understanding and core competencies that characterize accomplished interdisciplinary work.

Defining and Assessing Interdisciplinary Understanding

The term "interdisciplinary" is used in the literature to refer to a variety of intellectual activities ranging from the work of an individual scholar that bridges two of more domains to the work of collaborative teams, to fields emerging at the intersection of existing disciplines or problem centered work (Klein, 1996; Kockelmans, 1979; Newell, 1998; Rhoten & Pfirman, 2007; Weingart & Stehr, 2000). Common to these definitions is an emphasis on the synthesis or integration of information, perspectives, or insights stemming from different domains of knowledge. In this paper we define *interdisciplinary understanding* as the capacity to integrate knowledge and modes of thinking in two or more disciplines or established areas of expertise to produce a cognitive advancement—such as explaining a phenomenon, solving a problem, or creating a product—in ways that would have been impossible or unlikely through single disciplinary means (Boix Mansilla, Miller, & Gardner, 2000). "Understanding" is viewed here in a "performance" sense as the capacity to *use* knowledge flexibly and effectively, rather than *having* or *accumulating* it (Perkins, 1998). Assessing understanding

therefore requires that we give students opportunities to apply or *think with* the knowledge learned—opportunities that simultaneously build and demonstrate understanding. Instructors diagnostically examine students' work to support further development in understanding.

Building on this definition, the *Targeted Assessment Framework* (Boix Mansilla & Dawes Duraisingh, 2007) was designed to address long-standing questions of interdisciplinary instruction: What does it mean to understand a problem in interdisciplinary depth? What counts as quality interdisciplinary work? How can one describe core qualities of good interdisciplinary work to students in order to support their learning? The framework serves as a common language for describing a variety of interdisciplinary student products (multimedia presentations, papers, theses) highlighting core dimensions of interdisciplinary understanding applicable across cases.

First, the framework depicts quality interdisciplinary understanding as *grounded in disciplinary understanding*. Understanding builds on, expands, and revises knowledge and modes of thinking that have survived the scrutiny of expert communities using commonly agreed upon methods and validation standards (Gardner & Boix Mansilla, 1994). Psychological research in domain-specific cognition has documented challenges that students confront when developing disciplinary expertise such as overcoming intuitive misconceptions, linear causal reasoning, or social stereotypes. Students must also move beyond the view of disciplines as fixed bodies of information and understand the constructed and dynamic nature of knowledge production (Boix Mansilla & Gardner, 1998; Gardner 2000). Thus assessing interdisciplinary work involves a close *disciplinary* reading of student work, one that considers the foundational bodies of expertise on which a piece stands and the degree to which a student shows understanding of the chosen disciplinary insights, modes of thinking, or perspective—that is, preferred concepts, units of analysis, methods, and forms of communication in a discipline. Clearly, the inclusion of disciplinary insights or perspectives must be selective and driven by the purpose of the work in question.

Second, essential to interdisciplinary understanding is the *integration* of disciplinary insights and views that yields a new visual model, explanation, insight, or solution. Assessing the integrative qualities of a piece of work should therefore involve identifying points where insights from different disciplines have been brought together and articulating the cognitive advantage enabled by the combination of these insights. Third, the framework proposes that producing interdisciplinary understanding of quality demands that students have clarity about the purpose of their inquiry and engage in a process of considered judgment and critique:

weighing disciplinary options, making informed adjustments to achieve their proposed aims, recognizing the limitations of the work produced. Application of the *Targeted Assessment Framework* to practice revealed the need to distinguish between students' clarity of purpose and their capacity for self-critique. These two sub-dimensions are separated in our rubric.

A Grounded Rubric

Our *Targeted Assessment Rubric for Interdisciplinary Writing* includes 10 judgments grouped into four categories: (a) purposefulness, (b) disciplinary grounding, (c) integration, and (d) critical awareness or metacognition. On each of these dimensions, the papers are rated on a scale of 1-4 based on four levels: naïve, novice, apprentice, or master. In this section we characterize the four levels of understanding across dimensions and indicate the instructional effort that they invite. We then describe each dimension of interdisciplinary understanding, outlining the particular learning demands that they present to learners working on interdisciplinary papers.

Four Levels of Interdisciplinary Understanding

The levels of interdisciplinary understanding are presented as ideal types. They offer a coherent characterization of a piece of student work. In reality, however, a single student's paper may represent more than one level, depending on the dimensions assessed. For example, a student could produce a paper that shows a clear purpose and robust disciplinary grounding but only limited synthesis of disciplinary insights. Another student may articulate integrations well in a paper but exhibit disciplinary misconceptions that compromise its quality. Ultimately, this rubric seeks to diagnose student learning at a level of granularity that enables further instructional support.

Naïve interdisciplinary understanding. A paper can be characterized as naïve when it lacks clarity about purpose and audience; it is built primarily on common sense or folk beliefs about the topic at hand and the student fails to draw on disciplinary insights. Indeed, there is no effort to integrate multiple perspectives because perspectives themselves are not considered as such. Students who exhibit a naïve understanding of the problem under study will benefit from instruction that engages them in wondering about the topic in ways that may call for interdisciplinary exploration.

These students may benefit from discussions about why the topic matters, what would be gained by understanding it in depth and how to go

about developing this understanding, how the topic connects with and expands personal experience, or how their intuitions about the topic may be challenged.

Novice interdisciplinary understanding. A piece of work exemplifies a novice understanding when it exhibits a student's nascent grasp of the nature of interdisciplinary academic work. Often such work is mechanistic and tentative; it is informed primarily by the rituals of schooling and presents important composition weaknesses. At this level, the paper may have been undertaken in the spirit of writing an interdisciplinary paper *per se* rather than seeking to advance understanding of a complex problem. Disciplinary concepts and theories are typically presented as matters of fact, and misconceptions may be abundant. When novice writers do seek to incorporate multiple disciplines, the purpose of their paper may be too broad and unviable, or the integrative language may be mechanistic and pro forma. Students may balance perspectives according to superficial criteria, such as allocating equal length of text for each discipline.

Unlike their naive counterparts, novice writers are beginning to engage in academic writing and may be intrigued by a problem to be examined. These students will benefit from support in understanding the nature of and differences between disciplinary and interdisciplinary work and the process of inquiry. Students at this level may find it helpful to analyze examples of work by experts in which the process of knowledge construction in and across disciplines is made visible, as well as to reflect critically on the benefits and shortcomings of the different disciplinary contributions.

Apprentice interdisciplinary understanding. Papers at this level mirror experts' interdisciplinary work. Students exhibit a clear and viable purpose and a sense of the multiple audiences for the work. They adequately use disciplinary concepts and modes of thinking and support key claims with examples and sources. Integration is reached through a metaphor, conceptual framework, causal explanation, or other device that contributes to a deepening understanding of the topic. The paper may still include unnecessary diversions and opportunities to strengthen the argument may be missed. Students whose work falls primarily at this level are attaining a robust understanding of disciplinary foundations and how and why integration can deepen understanding of the topic at hand.

Further support for these students could involve polishing the piece to maximize its effectiveness and critical stance. Students may be invited to compare their interdisciplinary approach to the topic to competing scholarly ones to discern how understandings may be best advanced.

They may also be well served by in-depth reflection on the shifting terrain of existing disciplinary or knowledge structures and how their inquiry could be further advanced by deeper levels of integration.

Master interdisciplinary understanding. Papers at the master level are characterized by their creativity, parsimony, and sophisticated self-reflection. At this level, students demonstrate comfortable understanding of disciplinary foundations and interdisciplinary integration. Their papers exhibit a clear sense of purpose and need for an interdisciplinary approach. Students have mastered multiple expressive genres and can effectively create a hybrid form. They introduce new insightful examples to support disciplinary claims. Perspectives are integrated elegantly and coherently, and opportunities to advance the argument are not overlooked.

As undergraduates, students performing at this level are ready to move to a new topic. As graduates, new criteria can be considered, such as originality, potential impact of the work, and whether scholarly precedents and contributions have been accounted for. While a tested progression of interdisciplinary cognition is beyond the scope of this paper, the characterization we propose serves as a diagnostic tool to select appropriate interventions and personalize instruction.

Four Dimensions of Interdisciplinary Understanding

To introduce the four dimensions of our rubric, we begin with a brief description of a piece of student work on which we will draw for illustration.

Rebecca's Paper. "At the root of all modern political and societal organizations is a philosophical theory of human nature" (p. 1), claims Rebecca in the introduction to her final essay for a sophomore course on human nature. She asserts that differences between political and social systems can be attributed to their "different conceptions of human nature" (ibid). To explore this contention, Rebecca analyzes Marx's vision of the ideal classless society, taking the perspectives of sociobiology, capitalist economics, and sociology. She claims that for sociobiologists, "It is the biological nature of humans to be socially stratified. . . . A classless society seems to be impossible from their standpoint" (p 3). Rebecca later adds: "The conception of human nature based upon selfishness, individualism, and competition held by capitalist economists and game theorists does not allow for a [Marxist] society based upon community, cooperation and compassion."(p. 6). However, taking a sociological perspective, she concludes that Marx's utopia is theoretically possible because ideology has the power to create or approximate a new social reality: "When ideologies are internalized they become resistant to change and in fact a sort of reality for believers" (p. 7). She concludes

by putting forth her personal vision for an ideal society—one that is based on competition but which supports those who are biologically disadvantaged.

Employing the rubric to assess Rebecca's interdisciplinary understanding involves considering four fundamental aspects of her work: *purposefulness, disciplinary grounding, integration,* and *critical awareness.* In each case, the rubric invites us to gauge the quality of her performance and consider what her next learning challenge might be.

Purposefulness.

This dimension examines the degree to which students exhibit clarity about the aims and audience of their interdisciplinary writing. Two key questions guide assessment here:

Does the student's framing of the problem invite an integrative approach? Interdisciplinary work challenges students to present their object of study in ways that are both viable in scope and clearly multidimensional. In Rebecca's paper, the opening and concluding paragraphs suggest an ambitious intellectual goal: to discern whether philosophy is relevant in the modern world. However, the main body of her paper is concerned with the more manageable issue of examining whether Marx's ideal of a classless society is conceivable given the "theories of human nature" advanced by her three chosen scholarly traditions. The question about the theoretical feasibility of Marx's ideal is clearly stated and demands an integrative approach, placing her paper at an "apprentice" level. To move to a "master" level, Rebecca would be advised to restrict her paper's focus to this issue and to present a more explicit rationale for taking an approach which compares different disciplinary perspectives.

Does the student use the writing genre effectively to communicate with his or her intended audience? To produce a piece of interdisciplinary writing students must navigate domains that often embody contrasting discursive forms. Rebecca's paper uses appropriate tone and vocabulary. Some slips in register and several categorical claims place her work at an *apprentice* level—for example, "Philosophy today gets no respect,' says psychologist Steven Pinker, and he is dead right." Her most important discursive accomplishment is evident when she describes how readers operating in each of the traditions she reviewed would react to her synthetic position on classless society.

Disciplinary grounding.

This dimension examines students' understanding, selection, and use of the bodies of expertise that inform their work. Two questions guide assessment in this case:

Does the student use disciplinary knowledge accurately and effectively (e.g., concepts, theories, perspectives, findings, examples)? Grounding a paper on robust disciplinary understanding is an essential marker of quality work. A key conceptual change for students involves the shift from summarizing isolated bits of information from multiple disciplines to building flexible understanding of the disciplinary concepts to be borrowed. In masterful writing, students move effectively between abstract ideas and concrete examples, creating a rich conceptual network. In Rebecca's paper, concepts such as "natural selection," "class," and "conflict" were accurately employed and associated to not one but multiple experts in each domain. She cites relevant and credible sources such as works by E. O. Wilson and Steven Pinker. In most cases, Rebecca does not merely summarize expert definitions but applies them to the issue of the Marxist ideal of a classless society. With very few adjustments, Rebecca's paper could exhibit a master level in this dimension.

Does the student use disciplinary methods accurately and effectively (e.g., experimental design, philosophical argumentation, textual analysis)? Students' understanding of how knowledge is constructed in the fields relevant to their paper constitutes an important learning benchmark. Accomplished students recognize methodological differences and similarities in forms of knowledge production across domains and explore their tensions and complementarities. Rebecca's work is rooted primarily in philosophical argumentation. She prioritizes underlying assumptions and logical compatibility or incompatibility among intellectual positions. She alludes to the constructed nature of knowledge with comments such as "the sociological worldview seems to be the most difficult to argue against, but there is no absolute proof that it is correct." In a further iteration of this piece, Rebecca could consider the evidentiary forms employed by sociobiologists, economists, and sociologists. Such attention to how forms of knowledge production relate to claims and assumptions would considerably strengthen this already accomplished piece and move it from an apprentice to a master level.

Integration

Key to our definition of interdisciplinary understandings is students' capacity to *integrate* perspectives. Presented as a detailed characterization of the phenomenon of integration, this dimension examines how perspectives are selected, how connections across disciplines are framed, how they are articulated into a coherent whole, and what the advantage is of such articulation. Four criteria guide assessment at this level:

Does the student include selected disciplinary perspectives or insights from two or more disciplinary traditions that are relevant to the purpose of the paper? Clearly not all disciplines will prove relevant to the issue under study; neither will all theories, concepts, or modes of thinking available within a particular disciplinary tradition be applicable. Interdisciplinary work calls for purposeful disciplinary selection in terms of both which disciplines as well as which concepts and modes of thinking within them will inform the work. In her paper, Rebecca makes reasonable selections both of disciplines (philosophy, sociobiology, economics, and sociology) and particular insights within each discipline (natural selection, ideology, and classless society). These perspectives offer contrasting views of what underpins social contracts and in this regard, her work can be considered to be at master level.

Is there an integrative device or strategy (e.g., a model, metaphor, analogy)? At the heart of any interdisciplinary integration lies an integrative device—for example, a metaphor, complex explanation, or bridging concept—that brings together disciplinary insights. Occasionally, as in Rebecca's paper, integration is framed around a particular bridging concept—here, classless society—as it plays out in a variety of disciplinary contexts to yield a nuanced understanding. After concluding that Marx's ideal is incompatible with sociobiological and economist perspectives, Rebecca reaches a workable integration that borrows insights from each tradition: a society that recognizes its ties to human biology but overcomes undesired biological impulse through the power of culture and ideology. This is a master level performance.

Is there a sense of balance in the overall composition of the piece with regard to how the student brings the disciplinary perspectives or insights together to advance the purpose of the piece? This dimension examines the relative attention given to particular disciplines in the construction of an interdisciplinary argument. To reach an adequate balance, students must often overcome the inclination to give each discipline an equal share—a mechanistic approach common among novice interdisciplinary writers. Quality interdisciplinary papers may exhibit a dominant discipline if the purpose of the paper requires. Other disciplines may adequately play a "supporting role" as long as the concepts and modes of thinking in them are considered rigorously (as stated in Dimension 2 above). Disciplinary balance, as this guiding question suggests, is to be determined by the purpose of the piece at hand. On the whole, Rebecca gives appropriate weight to the different disciplinary lenses she brings to bear on the issue of classless society. Her comparative argument does require a comparable treatment of sociobiology, economics, and sociology—something she achieves at least in part, despite presenting the

sociobiological position with more nuance than the other two. However, as she reaches her conclusion, some of Rebecca's claims stand ungrounded and unexplained. For example, she argues that "more feminist-based characteristics could balance the aggressive male tendencies dominant in our societies," offering no further elaboration of this claim. With some adjustments, Rebecca could start to move toward a master level for this criterion.

Do the conclusions drawn by the student indicate that understanding has been advanced by the integration of disciplinary views? This question assesses the effectiveness of an interdisciplinary integration. It essentially asks "Was the effort worth it? Did it yield a new, richer, deeper, broader, or more nuanced understanding?" If the integrative device sets the stage for the articulation of disciplines (a metaphor, complex explanation, or conceptual bridge), this question assesses the outcome. Not uncommonly, students announce how disciplines will come together in a paper but fail to fulfill the promise. Rebecca's disciplinary integration yields key results: sociobiological and economic perspectives shed doubt on the viability of Marx's ideal but in turn, a sociological perspective suggests that a sociobiological stance should not be taken to an extreme. As a result, she develops a new, more nuanced, and multidimensional stance on the question of human nature and society that is approaching master-level:

> I believe the majority of human beings work best in a competitive environment . . . individuals have unique abilities and some are more gifted than others. Ideally a society would encourage all to reach their full potential but support those who are biologically unable to perform equally. Nature is not fair but society should be. There will inevitably be stratification but not as drastic and harmful as is currently. (p. 8)

Critical awareness.

Finally, the critical awareness dimension calls attention to students' capacity to take a meta-disciplinary perspective on their interdisciplinary work and reflect explicitly about the craft of weaving disciplines together. Two questions guide assessment:

Does the student show awareness of the limitations and benefits of the contributing disciplines and how the disciplines intertwine? As the previous categories suggest, interdisciplinary work requires a deliberate intertwining of disciplinary perspectives and an assessment of disciplinary insights for their potential contributions and limitations. Less experienced writers may provide a pro forma critique such as "More research is needed on this topic." More accomplished ones may weigh the merits and limitations of the selected disciplines in turn against alternative selections avail-

able. Rebecca's paper—like many of the papers we assessed—could be developed in this area. Although she names the disciplinary perspectives she uses, she only fleetingly refers to how each could potentially limit or advance her argument. To improve the paper, and move from a novice to an apprentice level, Rebecca could explicitly consider the limitations of the strictly biological or monetary emphasis of sociobiology and economics when a cooperative social model is considered.

Does the student exhibit self-reflection? Finally, academic writing is strengthened when authors are aware of the limitations of their work. Throughout her paper, Rebecca employs a suitably tentative tone in her presentation of her ideal society; however, she does not attempt to critique (or justify) the integrative approach she takes in the paper, placing her work between a novice and apprentice level. A more self-reflective stance toward her own endeavor, especially given the meta-disciplinary nature of the paper itself, would make her argument more sophisticated.

Constructing the Targeted Assessment Rubric for Interdisciplinary Writing

The rubric was developed through an iterative process of theory, validation, and revision. Using the *Targeted Assessment Framework* as a point of departure, we identified lower and higher levels of accomplishment. We reasoned that an empirically grounded rubric assessing interdisciplinary learning should show developmental differences between the interdisciplinary writing produced by upper and lower division undergraduate students of interdisciplinary studies, as well as differentiate between the writing of disciplinary and interdisciplinary seniors. Our initial rubric included 10 judgments grouped into four categories: (a) purposefulness; (b) disciplinary grounding; (c) integration; and (d) critical awareness. On each of these dimensions, the papers were rated on one of four levels: naïve, novice, apprentice, or master level.

Sampling

We collected 84 essays on a variety of topics in the humanities and social sciences that students had produced as part of their regular course assignments at Miami University. Sixty four essays were written by students in the School of Interdisciplinary Studies at Miami University who take a sequence of interdisciplinary courses preparing them for an individually tailored interdisciplinary major. Each course has learning outcomes that steadily increase in complexity, culminating in a year-long interdisciplinary thesis (Haynes, 2004), making essays from this population a particularly appropriate test of our theory and rubric.

The selected essays ranged in length from 4 to 115 total pages. Interdisciplinary freshman essays ($N = 24$) were typically three to five pages long. Interdisciplinary sophomore essays ($N = 20$) were 10–15 pages long. Interdisciplinary senior projects ($N = 20$) were 80–90 pages long. To offer a point of comparison, 20 disciplinary senior honors projects ($N = 20$) outside of the interdisciplinary studies programs were selected as well. These were typically 50–60 pages long. All students consented to have their essays used for this research by responding to course specific e-mail requests. Student names and all identifying information were removed from the essays by an assistant not directly connected to this project, and essays were scored only by a code number.

Rubric Development and Reliability Test

Groups of essays ($N = 10$, $N = 10$ and $N = 20$) were randomly selected from the set of 84. Each set was independently scored by two of the authors of this paper at Miami University who settled discrepancies by consensus. The other two authors at Harvard Graduate School of Education independently scored or shadowed the scoring of each set to further calibrate the rubric. After each round of scoring, the rubric was adjusted to reflect agreements and judges re-scored essays to reflect rubric modifications. The two Miami judges independently scored the remaining 44 essays. Inter Rater Reliability (IRR) was assessed by the number of judgments on which the four raters exactly agreed divided by the total number of judgments: IRR = 83.5%.

Validity Test

The validity of the rubric was tested with a series of four 4 x 1 ANOVAs corresponding to each dimension of the rubric. We hypothesized that interdisciplinary seniors would score significantly higher than interdisciplinary first-year and sophomore students on all categories. For each category our hypotheses were further tested with Fisher's PLSD least significant difference tests to distinguish among specific groups. Our hypotheses were confirmed (see Table 1). Senior interdisciplinary student essays scored significantly higher than their freshman and sophomore counterparts on the four rubric dimensions. On the ability to frame the *purpose* of their paper so as to invite an interdisciplinary approach, $F(3,80) = 15.86$, $p < 0.0001$; Fisher's PLSD = 0.327, $p < 0.0001$. On the ability to employ *disciplinary* concepts modes of thinking adequately, $F(3,80) = 38.80$, $p < 0.0001$, Fisher's PLSD = 0.323, $p < 0.0001$. On the exhibited capacity to *integrate perspectives* $F(3,80) = 14.717$, $p < 0.0001$, Fisher's PLSD = 0.421, $p < 0.0001$. Finally, on the

ability to offer a *self critical and meta-cognitive* perspective on their work, $F(3,80) = 19.058$, $p < 0.0001$, Fisher's PLSD = 0.346, $p < 0.0001$.

To further test the validity of our rubric, we applied it to a set of 20 senior *disciplinary* honors papers. The rubric is not intended for use with disciplinary papers. However, disciplinary senior honors essays were selected as a comparison group to test whether the rubric assesses more than generic writing quality. We reasoned that good disciplinary and interdisciplinary writing are grounded in disciplines but only interdisciplinary papers exhibit integration across disciplines. Therefore we predicted no differences between disciplinary and interdisciplinary senior papers regarding disciplinary grounding but significant ones regarding integration. We also understood that quality disciplinary and interdisciplinary papers have a clear purpose and exhibit a critical stance. However, we expected interdisciplinary papers to score slightly higher than their disciplinary counterparts due to the emphasis on interdisciplinarity at higher levels in the rubric. These predictions were confirmed (see Table 1).

Post Hoc Tests for Length of Essay

Not surprisingly, as the interdisciplinary students advance from freshman to sophomore to senior years, their essays become considerably longer. Considering pages of text excluding the title, tables, figures, and references, the essays were a mean of 3.9 (*SD* = 0.80) pages in the first year, 13 (*SD* = 3.46) pages in the sophomore year, and 83 (*SD* = 19.46)

TABLE 1

Mean Scores by Category and Group

Group	ID First Year (n = 24)	ID Sophomore (n = 20)	ID Senior (n = 20)	D Senior (n = 20)	Grand Mean (n = 84)
Purposefulness	2.229 (0.390)	2.175 (0.674)	3.100[a] (0.553)	2.825 (0.438)	2.565 (0.644)
Disciplinary Grounding	1.833 (0.381)	2.025 (0.550)	3.025[a] (0.743)	3.300 (0.441)	2.512 (0.825)
Integration	1.354 (0.321)	2.175 (0.487)	2.700[a,b] (0.724)	1.750 (1.082)	1.964 (0.855)
Critical Awareness	1.146 (0.275)	1.875 (0.582)	2.450[a,b] (0.776)	1.800 (0.594)	1.786 (0.737)
Grand Mean	1.641 (0.233)	2.062 (0.514)	2.819 (0.548)	2.419 (0.433)	2.207 (0.621)

Note. *SD* in parentheses. ID = interdisciplinary; D = disciplinary.

[a]Senior interdisciplinary project mean was significantly higher than sophomore or freshman interdisciplinary project means, $p < 0.0001$. [b]Senior interdisciplinary project mean was significantly higher than disciplinary senior project mean, $p < 0.001$.

pages of text in the senior year of interdisciplinary study. Indeed there was a positive correlation between the number of pages and each of the four outcome variables, for (a) purposeful understanding and pages of text, $r(62) = 0.62$, $p < 0.0001$; for (b) grounded in disciplines and pages of text, $r(62) = 0.72$, $p < 0.0001$; for (c) integration and pages of text, $r(62) = 0.42$, $p < 0.0001$; and for (d) critical awareness and pages of text, $r(62) = 0.53$, $p < 0.0001$.

To address the issue of page length as a potentially confounding variable, we conducted a set of four *post hoc* ANCOVA tests using pages of text as the covariate and group (freshman, sophomore, or senior year of interdisciplinary study) as the independent variable. The logic behind these tests is that if page length were the only important variable then group (class year) would no longer be significant after controlling for page length. On each of the four outcome measures, group significantly predicted outcomes. For (a) purposeful understanding, group was significant at $F(2,58) = 8.59$, $p = 0.001$. For (b) grounded in disciplines, group was significant at $F(2,58) = 4.13$, $p < 0.02$. For (c) integration, group was significant at $F(2,58) = 4.48$, $p < 0.02$. Finally, for (d) critical awareness, group was significant at $F(2,58) = 5.84$, $p < 0.005$. Further testing of interaction effects between group and page length confirmed that page length was not predictive of performance in all but two cases where page length predicted outcomes within one class level but not the other two. In sum, these results suggest that developmental differences in the quality of interdisciplinary writing found between first year and senior students are a function of educational experience rather than solely the number of pages they wrote. Despite the high correlation between class year and essay length, class year remained significant on each outcome variable, even controlling for pages of text. These results indicate that the rubric can be used reliably to score undergraduate essays covering a diverse range of topics, as well as validly to capture developmental differences among interdisciplinary studies students. In our conclusion we turn our attention to the adaptation and use of the rubric.

Conclusion

Disputes about definitions of interdisciplinary work, the broad variety of disciplinary combinations possible, and the simple fact that interdisciplinary understanding is a complex cognitive enterprise, have militated against clear indicators of quality interdisciplinary student learning. By proposing a theoretically grounded and empirically tested rubric to assess student interdisciplinary writing, we seek to shed new light on the issue of interdisciplinary assessment and research on integrative

thinking. The rubric outlines an anatomy of interdisciplinary thought and its progression, offering multiple potential applications. All applications will demand a calibration of the rubric descriptors to ensure appropriate use.

In instructional contexts, the *Targeted Assessment Rubric for Interdisciplinary Writing* may be shared with students, ideally with the descriptors reframed to capture the content of the course. For instance, if history and physics are considered as key disciplines in a given course, the particular theories, historical narratives, authors, and approaches could be included under "disciplinary grounding." If a particular form of integration is desired, relevant descriptors could be adjusted accordingly. Applying the rubric to assess student work will enable faculty and students to know exactly what has been accomplished and what else could be done to advance the work. The rubric can also inform student peer and self-assessment.

This rubric has been tested on academic papers. However, with minor adjustments, it could be applied to a variety of forms of student work. In fact, the *Targeted Assessment Framework* on which this instrument is based was created from close analysis of a variety of student learning outcomes such as written papers, presentations, and works of art accompanied by written reflections. In all cases, categories and levels remain relevant, but particular descriptors may be adjusted to open room for non-verbal and non-written modalities.

By focusing on interdisciplinary student learning, the rubric sets the foundation for a valid and rigorous evaluation of interdisciplinary programs—one that prioritizes student performance over self reports and targeted assessment of interdisciplinary capabilities over more generic "grades." Furthermore, the rubric enables us to compare the performance of students at different moments in a multi-year interdisciplinary program. Although the rubric is not designed to assess *disciplinary* papers or performances per se, a valid comparison between graduates of disciplinary and interdisciplinary training programs can be made on the basis of a common interdisciplinary performance task. In fact, that was the choice made by Rhoten and Hackett in their evaluative study of NSF's IGERT Program. In this case, researchers and disciplinary experts adapted the rubric to assess the performance of beginning and graduating students in interdisciplinary and disciplinary doctoral programs in environmental sciences (Haag, 2006).

Rooted in the *Targeted Assessment Framework's* principles of cognition and epistemology and tested on interdisciplinary papers, this rubric offers a preliminary framework for studying learning progressions in interdisciplinary thinking. Researchers interested in this area of work may

seek to adapt and apply such a framework in a longitudinal study, complemented with empirical findings in progressions in disciplinary thinking, well-tested measures of epistemological beliefs, and adequate controls for learning style and disciplinary selection. A cognitive developmental portrait of progressions in interdisciplinary thinking will greatly inform the identification of benchmarks and misconceptions associated with interdisciplinary learning. The *Targeted Assessment Framework* and *Rubric* should inform a more rigorous and systematic study of interdisciplinary cognition—a phenomenon on which the greatest technical, social, and environmental advancements of the century will rely.

References

Astin, A. W. (1993). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education*. Phoenix, AZ: Oryx Press.

Banta, T. W. (Ed.). (2002). *Building a scholarship of assessment.* San Francisco: Jossey-Bass.

Boix Mansilla, V. (2005). Assessing student work at disciplinary crossroads. *Change* 37 (January/February), 14–21.

Boix Mansilla, V., & Dawes Duraisingh, E. (2007). Targeted assessment of students' interdisciplinary work: An empirically grounded framework proposed. *The Journal of Higher Education 78*(2), 215–237.

Boix Mansilla, V., & Gardner, H. (1998). What are the qualities of disciplinary understanding? In M. S. Wiske (Ed.), *Teaching for understanding: Linking research with practice* (pp. 161–196). San Francisco: Jossey-Bass.

Boix Mansilla, V., Miller, W. C., & Gardner, H. (2000). On disciplinary lenses and interdisciplinary work. In P. Grossman & S. Wineburg (Eds.), *Interdisciplinary curriculum: challenges to implementation* (pp. 17–38). New York: Teachers College Press.

Brough, J. A., & Pool, J. E. (2005). Integrating learning and assessment: The development of an assessment culture. In J. Etim (Ed.) *Curriculum integration K-12: Theory and practice* (pp. 196–204). Lanham, MD: University Press of America.

Chun, C. (2002). Looking where the light is better: A review of the literature on assessing higher education quality. *Peer Review,* Winter/Spring, 16–25. Retrieved June 13, 2007, from http://www.aacu.org/peerreview/pr-sp02/pr-sp02feature3.cfm

Cizek, G. J. (2000). Pockets of resistance in the assessment revolution. *Educational Measurement: Issues and Practices, 19*(2), 16–23.

Ewell, P. T. (1991). To capture the ineffable: New forms of assessment in higher education. *Review of Research in Education,* (17), 75–125.

Gardner, H. (2000). *The disciplined mind: Beyond facts and standardized tests, the K-12 education that every child deserves.* New York: Penguin Books.

Gardner, H., & Boix Mansilla, V. (1994). Teaching for understanding in the disciplines and beyond. *Teacher College Record, 96*(2), 198–218.

Goodrich-Andrade, H. (2006). The trouble with a narrow view of rubrics. *English Journal, 95*(6), 9.

Haag, A. (2006, September). A Testing Experience. *Nature* 443, 265–267.

Haynes, C. (2004). Promoting self-authorship through an interdisciplinary writing curriculum. In M. B. Baxter Magolda & P. M. King (Eds.) *Learning partnerships: Theory and models of practice to educate for self-authorship* (pp. 63–90). Sterling, VA: Stylus.

Huba, M. E., & Freed, J. E. (2000). *Learner centered assessment on college campuses: Shifting the focus from teaching to learning.* Needham Heights, MA: Allyn & Bacon.

Huber, M. T., & Hutchings, P. (2004). *Integrative learning: Mapping the terrain*. Washington: Association of American Colleges and Universities.

Hutchings, P. T. (1990). Learning over time: Portfolio assessment. *American Association of Higher Education Bulletin, 42*, 6–8.

Klein, J. T. (1996). *Crossing boundaries: Knowledge, disciplinarities, and interdisciplinarities*. Charlottesville, VA and London: University Press of Virginia.

Kockelmans, J. J. (1979). Why interdisciplinarity? In Kockelmans (Ed.) *Interdisciplinarity and Higher Education*. University Park, PA: Pennsylvania State University Press.

Kohn, A. (2006). The trouble with rubrics. *English Journal, 95*(4).

Levy, F., & Murnane, R. (2004). *The new division of labor: How computers are creating the next job market.* Princeton, NJ: Princeton University Press; New York: Russell Sage Foundation.

National Academy of Sciences, National Academy of Engineering, and Institute of Medicine of the National Academies (2005). *Facilitating interdisciplinary research.* Washington, DC: The National Academies Press.

National Institutes of Health (2006). Research teams of the future. In *NIH Roadmap for Medical Research.* Retrieved June 17, 2007, from http://nihroadmap.nih.gov/researchteams/

National Leadership Council for Liberal Education and America's Promise (2007). *College learning for the new global century. Executive summary.* Washington, DC: Association for American Colleges and Universities.

National Science Foundation (2006, September). *Investing in America's future: Strategic Plan FY 2006-2011.* Retrieved June 13, 2007, from http://www.nsf.gov/pubs/2006/nsf0648/NSF-06-48.pdf

Newell, W. H. (1998) Interdisciplinarity: Essays from the literature. New York: The College Board.

Perkins, D. (1998). What is understanding? In M. S. Wiske (Ed.), *Teaching for understanding: Linking research with practice* (pp. 39–57). San Francisco: Jossey-Bass.

Rhoten, D., Boix Mansilla, V., Chun, M., & Klein, J. T. (2006). *Interdisciplinary education at liberal arts institutions.* Teagle Foundation White Paper. Retrieved June 13, 2007, from the Teagle Foundation Web site: http://www.teaglefoundation.org/learning/pdf/2006_ssrc_whitepaper.pdf

Rhoten, D. & Pfirman, S. (2007). Women in interdisciplinary science: Exploring preferences and consequences. *Research Policy, 36*, 56–75.

Schilling, K. L. (2001). Interdisciplinary assessment for interdisciplinary programs. In B. L. Smith & J. McCann (Eds.), *Reinventing ourselves: Interdisciplinary education, collaborative learning and experimentation in higher education* (pp. 344–354). Bolton, MA: Anker Publishing Company.

Schneider, C. G. (2002). Can value added assessment raise the level of student accomplishment? *Peer Review, 4*(2/3), Winter/Spring.

Walvoord, B. E. F., & Anderson, V. J. (1998). *Effective grading: A tool for learning and assessment*. San Francisco: Jossey-Bass.

Weingart, P. & Stehr, N. (2000). *Practicing Interdisciplinarity*. Toronto, Ontario, Canada: University of Toronto Press.

Wilson, M. (2006). *Rethinking rubrics in writing assessment.* Portsmouth, NH: Heinemann.

Wolfe, C.R., & Haynes, C. (2003). Interdisciplinary writing assessment profiles. *Issues in Integrative Studies*, *21*, 126–170.