
TOPICAL ARTICLES

A Rubric for Learning, Teaching, and Assessing Scientific Inquiry in Psychology

Jane S. Halonen
James Madison University

Ted Bosack
Providence College

Shirley Clay
Northeast Texas Community College

Maureen McCarthy
Austin Peay State University

In Collaboration With

Dana S. Dunn
Moravian College

G. William Hill IV
Kennesaw State University

Rob McEntarffer
Lincoln Southeast High School
Lincoln, Nebraska

Chandra Mehrotra
The College of St. Scholastica

Robbye Nesmith
Navarro College

Kenneth A. Weaver
Emporia State University

Kristin Whitlock
Davis County School District
Utah

Accountability pressures influence all levels of psychology instruction. In this article we explore how to meet those pressures with integrity, focusing on authentic assessment and teaching as a primary solution. We propose a rubric to describe the progress of students' acquisition of scientific inquiry skills applied to behavior and pro-

vide an example of an authentic assessment that demonstrates use of the rubric. Application of the rubric can enhance active learning, promote more sophisticated scientific inquiry, improve metacognitive development, support program evaluation, and enrich faculty development.

Teachers of psychology across all contexts—from high school through graduate institutions—increasingly feel challenged by external pressures to produce evidence regarding the quality of what they do. State boards of education and accreditors bring substantial pressures to program administrators for demonstrating improved performance (O’Neil, 1992). Although external pressures vary with educational contexts, nearly all teachers of psychology have responded to a broader range of constituents in the planning and execution of their courses as the result of the growth of assessment practices (Schneider & Schoenberg, 1998).

Ideally, assessment should promote competence in students (Mentowski et al., 2000; Wlodkowski & Ginsberg, 1995). In addition, assessment improves educational practice by offering benchmarks or standards of performance against which teachers can measure student learning (Astin, 1993; Banta, Lund, Black, & Oblander, 1996). The results of assessment, furthermore, should help stakeholders to make informed judgments about achievement and teaching quality. These results should also provide direction for methods to improve learning.

Selection of appropriate assessment strategies produces concern (Maki, 2001). Assessment pressures prompt administrators to use traditional objective testing strategies that may minimize more sophisticated or subtle aspects of the knowledge and abilities that students learn. Expedience can encourage the use of quantitative measures that necessarily tap less sophisticated indicators of learning. The use of standardized, quantitative measures can sometimes encourage teachers to misdirect their energies by attempting to “teach to the test.” They cope with the measures that are convenient or expedient rather than using other assessment strategies that facilitate greater student involvement. As a consequence, some teachers strive to appear to use effective assessment rather than truly develop effective assessments that will improve learning and teaching.

Although quantitative measures of student progress are useful and necessary, authentic assessment techniques (Palomba & Banta, 1999) offer additional measures of student performance. Authentic assessment employs real-world tasks, asking students to engage in meaningful activities, thereby providing task relevance to the students and allowing for evaluation of effective performance. Furthermore, authentic assessments allow for evaluation of the entire process required for completion of a product, thus offering a stronger basis for validity (Wiggins, 1990). Despite the advantages of using qualitative measures, implementation remains problematic (Strong & Sexton, 1996). For example, critics challenge use of student portfolios on the basis of cost and storage of student materials, time-intensive feedback responsibilities, and judgment fairness and reliability. Notwithstanding these problems, Strong and Sexton (1996) reported that authentic assessment strategies are flourishing.

In this collaborative article, we propose a developmental rubric to assist psychology teachers across educational contexts to assess scientific inquiry from a systematic qualitative perspective. In essence, this model responds to concerns of evaluators and integrates cognitive development research into the evaluation process to specify outcomes in behavioral terms. The rubric articulates developmental expectations about how we think students’ skills typically progress over the

course of their education in psychology. Acceptance of the rubric has clear implications for assignment design, program evaluation, and faculty development. As such, the rubric can assist with the variety of assessment challenges that face psychology teachers and administrators.

Background of the Rubric

Formulation of this rubric began at the Psychology Partnerships Project (P3), an initiative sponsored by the American Psychological Association (APA; Andreoli Mathie & Ernst, 1999). P3 grew out of separate requests to the APA to convene curriculum-focused conferences at the high school and undergraduate levels. Virginia Andreoli Mathie, representing 4-year college concerns, and Randall Ernst, representing the Teachers of Psychology in Secondary Schools, combined their proposals to the Board of Educational Affairs of the APA and added the theme of promoting partnerships across educational contexts to solve common problems. In 1997, the Education Directorate assembled a Steering Committee that involved teachers from all levels of education in psychology (i.e., high school, 2-year or community college, college and university, graduate education). The Steering Committee designed a National Forum in which participants worked in groups to collaborate on how common issues (e.g., technology, faculty development, service learning) could be explored through partnerships across contexts. In June 1999, 90 educators representing each educational level or context convened at James Madison University for the 5-day forum.

The Assessment Group, under the leadership of Rob McEntarffer (Lincoln Southeast High School), focused on how to make assessment strategies more effective and supportive of the training that psychologists do in the classroom. Ultimately the group settled on two ideas to improve assessment practices in psychology: the development of a rubric to promote authentic assessment of scientific inquiry and a conference to promote the best practices of educational assessment in psychology, which took place in September 2002.

Development of the rubric centered on Halpern’s (1988) suggestion that solid outcome measures determine the “value added” from the education experience. We embraced her ideas by trying to capture qualitatively the learning at each level of psychology education from the standpoint of scientific psychology. We identified specific dimensions of scientific inquiry as they apply to describing, explaining, controlling, and predicting behavior, and we reviewed how those dimensions emerge during education in psychology to propose qualitative descriptors of student change. Eventually, our discussions produced a rubric, a system for assessing complex student responses (Trice, 2000), on the development of scientific inquiry. Group members then shared the rubric in various public settings and with colleagues to improve its utility and language and to verify their speculations about student progress across time (e.g., Dunn & McCarthy, 2001).

We propose that this rubric promotes authentic assessment strategies to facilitate the opportunity for capturing the more qualitative aspects of meaningful and enduring learning. Attending to these qualitative benchmarks encourages the development of objective standards or criteria to evaluate

achievement of these benchmarks. Adoption of the rubric provides a framework to promote the relevance of course experience to students' lives, and, as such, the rubric becomes a road map for teaching effectiveness.

A Developmental Rubric for Scientific Inquiry

Defining Scientific Inquiry Domains

Morgan and Johnson (1997) asserted that psychology has no central paradigm that can be tested, which complicates the establishment of benchmarks for psychology education (cf. Kuhn, 1962; Watson, 1967). Curriculum scholars (e.g., Brewer et al., 1993; McGovern, Furumoto, Halpern, Kimble, & McKeachie, 1991), however, converged on the importance of instilling scientific reasoning skills through studying a diverse set of content areas as a fundamental objective of psychology. In an effort to relate the study of scientific psychology to authentic assessment, Hogan (1999) advocated that students should learn to read about and accept scientific principles but should also learn how to construct and evaluate scientific principles as part of the general goals of training in science. Zimmerman (2000) recommended that science students should have opportunities to form hypotheses, design experiments, observe, measure, make inferences, and formulate theories. Furthermore, students should engage in problem-solving strategies that grow out of the discipline's theories.

Consistent with these perspectives, the Assessment Group of P3 formulated developmental standards in the context of psychology that reflect the ability to think about and identify psychological problems from an empirical perspective. The resultant outcome was the rubric for scientific inquiry in psychology that incorporated the objectives of describing, explaining, controlling, and predicting behavior. Hogan and Fisherkeller (2000) originally offered four phases of scientific

inquiry that included (a) forming hypotheses, (b) designing experiments, (c) interpreting outcomes, and (d) communicating results. We expanded these phases and articulated eight domains of scientific inquiry that could serve as outcomes for scientific inquiry in psychology. We further subdivided each domain into specific skill areas and listed these abilities in order of their emergence or importance. The domains and skill areas include the following:

1. Descriptive skills: observation, interpretation, and measurement skills (see Table 1).
2. Conceptualization skills (the ability to use the concepts and theories of the discipline): recognition and application of concepts, recognition and application of theory, and advanced theory skills (evaluation, synthesis, and generation; see Table 2).
3. Problem-solving skills (the ability to conduct research and use research findings): recognition, evaluation, and generation of research methods; bias detection and management; and statistical reasoning (see Table 3).
4. Ethical reasoning: awareness, adherence (see Table 4).
5. Scientific values and attitudes: enthusiasm for research, objectivity/subjectivity, parsimony, skepticism, and tolerance of ambiguity (see Table 5).
6. Communication skills: resource-gathering skills, argumentation skills, and conventional expression (see Table 6).
7. Collaboration skills: projection completion skills, process management, consensus-building skills, leadership, and brainstorming (see Table 7).
8. Self-assessment: self-regulation, self-reflection (see Table 8).

Although some may choose to define scientific inquiry more narrowly, we have chosen this broader perspective, including skills in communication, collaboration, and

Table 1. Description of Proficiency Levels of the Descriptive Skills Domain

Components of Descriptive Skills	Levels of Proficiency				
	Before Training	Basic Introductory Psychology	Developing	Integrating Advanced Undergraduate	Professional Graduate and Beyond
Observation	Observes behavior superficially	Observes general patterns; tends to confuse observation and interpretation of behavior	Observes broadly and carefully; distinguishes observation from interpretation	Makes more subtle, sensitive observations that are distinct from interpretations	Uses sophisticated observational techniques appropriate to the circumstances
Measurement	Uses measurements in a limited or imprecise manner	Takes direction to measure critical elements but may resist demands for precision	Uses measurement as empirical strategy; complies with demands for precision	Actively pursues appropriate instrumentation; implements reasoned measurement strategy	Strives to produce optimal measurement strategy to reduce challenges to validity of conclusions
Interpretation	Relies on intuition that tends to focus on obvious and simplistic conclusions	Overinterprets behavioral events; forms judgment with limited regard to quality of evidence	Selectively combines data-based interpretations with personal experience	Relies on data more systematically for interpretation; produces more complex interpretations	Interprets behavior at appropriate level of complexity that reflects contextual factors

Table 2. Description of Proficiency Levels in the Conceptualization Skills Domain

Components of Conceptualization Skills	Levels of Proficiency				
	Before Training	Basic Introductory Psychology	Developing	Integrating Advanced Undergraduate	Professional Graduate and Beyond
Concept Skills: Recognition and application	Explains and predicts behavior without reference to scientific concepts; tends to rely on pop psychology interpretation	Recognizes connection between psychological concepts and observed behavior	Can identify and apply concepts purposefully to explain a behavioral event when prompted	Identifies and applies concepts independently to explain and predict behavior	Uses psychological concepts fluently to explain and predict behavior; evaluates the validity of concepts as explanations; generates new concepts
Basic theory skills: Recognition and application	Does not distinguish psychological theory from common sense	Can recognize theory elements in examples with guidance	Applies theory to explain and predict behavior but struggles when coping with counterintuitive conclusions from theory	Applies theory to explain and predict behavior and tolerates counterintuitive findings	Exploits discrepancy between intuitive findings and research as opportunity to explore new ideas
Advanced theory skills: Evaluation, synthesis, and generation	Does not recognize or favor the use of psychological theory	Evaluates theory quality based on its goodness of fit with personal beliefs	Identifies some objective constraints in theory use; Tests generalization of ideas by applying theory to new situations	Evaluates quality and fit of theory application objectively; compares and contrasts relative contributions of theories; integrates theoretical perspectives	Generates original theoretical explanations; Assumes responsibility for criticizing and improving theory

Table 3. Description of Proficiency Levels in the Problem Solving Skills Domain

Components of Problem Solving Skills	Levels of Proficiency				
	Before Training	Basic Introductory Psychology	Developing	Integrating Advanced Undergraduate	Professional Graduate and Beyond
Methods skills: Recognition, evaluation, generation	Does not rely on scientific method	Recites steps in conducting research; articulates basic knowledge of correlational and causal techniques; acknowledges value of controlled comparisons	Selects and applies appropriate method in simple projects; operationalizes and isolates variables; identifies influence of extraneous variables	Selects and applies appropriate method to maximize validity and reduce alternative explanations	Develops unique applications of research methods; establishes a research focus that identifies and build on primary interests in behavior
Statistical reasoning: Recognition, application, evaluation, and generation	Tends to use the mathematical term <i>average</i> improperly, ignoring its literal meaning	Uses basic descriptive statistics; accepts inferential analysis without understanding statistical foundations	Selects and applies appropriate statistical processes to simple projects; evaluates success of projects in global sense	Selects and applies appropriate statistics with more independence; begins to use statistical reasoning as a basis for criticizing research results	Uses statistical reasoning routinely for evaluating research; develops unique applications of statistics
Bias detection and management	Shows limited sensitivity to existence or effects of bias	Can recognize some potential sources of bias when prompted	Begins to recognize influence of bias and confounds in framing research questions	Identifies the potential influences of bias and confounds; recognizes personal, professional, and cultural interests influence problem and method selection	Implements and demands high standards for adherence to scientific method to minimize complications of bias and confounds

Table 4. Description of Proficiency Levels of the Ethical Reasoning Domain

Components of Ethical Reasoning	Levels of Proficiency				
	Before Training	Basic Introductory Psychology	Developing	Integrating Advanced Undergraduate	Professional Graduate and Beyond
Awareness of ethical standards	Shows limited awareness of or misconstrues general ethical practices in psychology	Recognizes existence of and rationale for ethical standards	Can identify how ethical standards apply to a given research example	Can identify how ethical standards enhance or constrain research	Monitors ethical practices in areas of research specialization
Evaluation of ethical practices	May assume that psychologists generally tend to be ethically misguided	Can identify gross violations of ethical standards in practice	Can apply ethical standards to given examples to judge the quality of ethical practice	Recognizes more subtle ethical breaches and suggests alternatives	Routinely evaluates research from an ethical standpoint as an ongoing professional responsibility
Adherence to ethical standards	Not applicable	Accepts ethical conditions required to participate in science but tends to question necessity	Accepts and adheres to prescribed ethical protocols when prompted	Executes appropriate ethical safeguards as a researcher's responsibility	Advocates for the best ethical practices to protect the public and improve the discipline

Table 5. Description of Proficiency Levels of the Scientific Attitudes and Values Domain

Components of Scientific Attitudes and Values	Levels of Proficiency				
	Before Training	Basic Introductory Psychology	Developing	Integrating Advanced Undergraduate	Professional Graduate and Beyond
Enthusiasm for research	Does not think about research findings	Accepts research findings that confirm personal experience; rejects disconfirming results	Allows selected scientific results to clarify personal experience	Views scientific method as valuable for personal and professional discovery	Practices scientific method and accepts its limitations
Objectivity/subjectivity	Relies solely on subjective/personal reality	Recognizes the difference between objective and subjective realities in limited, controlled, directed circumstances	Distinguishes objective and subjective reality; recognizes that perceptions of reality vary between individuals	Shows greater vigilance about avoiding negative outcomes of subjective influences	Guards against subjective influences
Parsimony	Shows little or no preference for a given explanation from a range of explanations	Recognizes that precise explanations tend to be better than imprecise ones	Attends to precision as an important aspect of research design	Uses precision as a criterion to determine quality of explanation	Seeks the most precise explanation
Skepticism	Accepts or fails to question the status quo	Tends to resist asking questions to avoid making situations too messy or too complex	Practices limited skepticism as externally, hedonistically driven (e.g., it's what the professor wants)	Practices skepticism selectively to improve evaluation skills	Uses skepticism consistently as an evaluative tool
Tolerance of ambiguity	Expects and accepts simple or easy behavioral explanations	Resists, protests, ambiguity	Begins to seek clarity in understanding behavior, but prefers simple/easy explanations over complex ones	Assumes behavioral explanations will be complex, and begins to tolerate ambiguity	Enjoys complexity during the search for clarifying behavioral explanations

Table 6. Description of Proficiency Levels of the Communication Skills Domain

Components of Communication Skills	Levels of Proficiency				
	Before Training	Basic Introductory Psychology	Developing	Integrating Advanced Undergraduate	Professional Graduate and Beyond
Resource-gathering skills: Selectivity, relevance, currency, quality of evidence	Relies on popular press reports of psychology and anecdotal evidence; selects sources based on personal bias	Reproduces simplistic, textbook capsules of information about behavior; conducts limited, expedient literature search; relies on secondary sources	Relates content from several sources, but tends to include resources nonselectively; shows limited use of formal literature	Integrates content from multiple sources, acknowledging contradictory information; uses resources selectively	Evaluates relevant content from broader range of available sources; reflects how context influences judgment; shows refined and flexible use of published research
Argumentation skills: Organization, awareness of audience, persuasiveness	Argues based on common sense; accepts personal experience as conclusive	Uses basic concepts to develop simple arguments; shows limited awareness of engaging audience; tends to argue from personal experience rather than from research evidence	Develops plausible arguments; demonstrates some awareness of audience by using engaging language and examples, but presumes audience knowledge is consistent with own	Creates coherent and integrated arguments based upon research evidence; engages audience by refined use of language, examples, and supports	Creates compelling arguments with attention to subtle meanings of content; anticipates and defends against criticism; adapts arguments for wide range of audiences
Conventional expression: Use of APA format, grammar, appropriate supportive visuals	Not relevant	Recognizes existence of APA format but questions its value and rigor; applies APA format with serious flaws; expresses ideas in informal language	Uses APA style inconsistently; shows increasing formality and professionalism in expression	Uses APA format more consistently; practices professional, formal expression to improve communication appeal	Uses APA format expertly with minimal errors; demonstrates sophisticated conventional expression

self-assessment, for a variety of reasons. First, by including communication skills among the domains of scientific inquiry, we concur in the assertion “research is complete only when the results are shared with the scientific community” (APA, 2001, p. 3). Second, the importance of collaboration skills emerges from an examination of representative sources for reporting scientific inquiry in psychology. Of the 51 research articles reported in *Developmental Psychology* in 2001, 49 (96%) had two or more authors, and 32 (63%) included three or more collaborators. Fifty of the 76 articles in *Teaching of Psychology* in 2001, representing 66% of the journal’s content, were collaborative efforts. At the 2002 convention of the American Psychological Society, 642 of the 745 posters presented (84%) carried multiple authorship. These figures reflect the magnitude of collaborative inquiry in psychology and the relevance of acquiring collaborative skills. Finally, in his discussion of the characteristics of science and the crafting of a science of behavior, Skinner (1965) observed that “science is a set of attitudes” (p. 12), among which is honesty in the evaluation of data and one’s scientific inquiry. Our inclusion of the domain of self-assessment affirms the importance of this attitude in the development of mature psychological scientists.

Development Over Time

We conceptualize the rubric as having three major developmental levels, beginning with the characteristics of the person who has not had any training in psychology (designated in the rubric as “Before Training”). To capture the characteristics of this level, we identified various scenarios to extract the fundamental features of untrained reasoning about behavior. For example, we recalled episodes of beginning students who were disappointed when we could not answer psychological questions with simple answers and construed what that disappointment might mean for their cognitive development. We applied this process to the other levels of the rubric.

The second level in the rubric—“Basic” through “Integrating”—captures the changes we might expect to see in students studying psychology from their earliest exposure to psychology (i.e., high school or college) through graduation as a bachelor’s-level psychology major. Three sublevels within this grouping distinguish the performance of students who have just completed their first introductory experience through those who have completed their degree. In effect, the second level of the rubric has three markers that may be

Table 7. Description of Proficiency Levels of the Collaboration Skills Domain

Levels of Proficiency					
Components of Collaboration Skills	Before Training	Basic Introductory Psychology	Developing	Integrating Advanced Undergraduate	Professional Graduate and Beyond
Project completion skills	Not relevant	Can complete simple projects with direction	Completes more complex projects with reduced direction	Completes more complex projects that may require collaboration over time with minimal direction	Independently completes sophisticated group projects that may require collaboration over time
Process management		Adheres closely to directions provided by authority figure; tends not to attend to quality of group process	Expects group members to contribute equal work to accomplish goal; begins to monitor group process to manage process more effectively	Collaborates to assign roles and responsibilities more strategically to achieve completion; evaluates quality of thinking produced by the group to improve process	Systematically plans project completion strategy, including back-up plans to overcome likely obstacles
Leadership		Expects and complies with leadership from appointed group leaders	Shares leadership or fulfills assigned responsibilities to help group achieve success	Exercises some leadership to contribute to positive working climate	Convenes colleagues to improve quality of projects and programs
Consensus building skills		Expects agreement; may be unable to function in conflict	Tends to stress value of own position ahead of others but recognizes that other positions may have merit	Can integrate diverse viewpoints to improve quality of group process and outcome	Conscientiously seeks expression of broad opinions and productive conflict resolution
Brainstorming		May not spontaneously generate creative alternatives in a structured situation	Can develop creative solutions/alternatives when encouraged	Contributes to building safe climate to reduce risk in generating creative alternatives	Actively promotes and enjoys group creativity

Table 8. Description of Proficiency Levels of the Self-Assessment Skills Domain

Levels of Proficiency					
Components of Self Assessment	Before Training	Basic Introductory Psychology	Developing	Integrating Advanced Undergraduate	Professional Graduate and Beyond
Self-regulation: Work completion skills	Works reactively in response to task demands without careful planning	Acknowledges role of planning to complete tasks appropriately but may not always be successful	Plans priorities reasonably to accomplish tasks	Executes complex projects properly, taking into account personal characteristics and challenges	Formulates back-up plans to anticipate and overcome obstacles
Self-reflection: Thinking about thinking	Tends not to reflect on own thinking or engage in self-reflection	Can make global self-assessments about quality of own work	Applies criteria to judge own performance with some consistency and accuracy; demonstrates some preference for shallow rather than deep critique	Makes refined, accurate judgments about quality of work; prefers deep criticism to improve quality of future work	Uses self assessments to establish goals

fairly fluid depending on the design of the curriculum and the nature of the students (see Table 1).

At the highest level, the "Professional" category describes intended outcomes of graduate education. We believe that these descriptors also apply to the scientific inquiry shown by professional practitioners, researchers, and educators. Tasks associated with the professional category possess higher levels of cognitive complexity (cf. Kitchener & Fischer, 1990) that reflect the final level of developed scientific inquiry. Assignments at this level usually reflect actual performance of professional psychologists. Similarly, the use of the rubric to evaluate their performance suggests a high level of integration rather than the demonstration of minimum competence levels.

We do not conceive of these developmental levels in the manner of Piaget's (1963) stages emerging from the interaction of maturation and experience, being universally ordered, and involving restructuring of cognitive processes. They do represent the acquisition of knowledge and understanding as a result of the teaching and learning process. The levels and behaviorally indexed qualities within them derive from lengthy analysis and collaboration (see Note 2). Consequently, we acknowledge the importance of providing an empirical foundation for the developmental progression and encourage scholars in the teaching of psychology to join this research effort.

Uses of the Rubric for Scientific Inquiry

It is easy for teachers, challenged by attending to the minutiae of a given course, to forget how their efforts will contribute to students' development over time. The rubric provides a good standard and meaningful context in which individual course objectives can be compared with how teachers expect students to develop.

A more subtle use of the rubric for understanding development is the recognition that students must develop through experience over time. For example, Bloom's taxonomy would suggest that learning begins with basic knowledge and proceeds developmentally through increasingly more complex levels (Linn & Gronlund, 1995). Teachers sometimes lose patience with students who ask simplistic questions about complex phenomena or who become agitated or unsettled when complex answers seem unsatisfying or (honestly) incomplete. The rubric reminds teachers that these behaviors are developmentally normal and challenges teachers to produce learning experiences that might help students evolve.

Designing Learning Experiences

Authentic assessment involves instructional design that promotes active learning (Barr & Tagg, 1995). Such assessments use realistic contexts to evaluate higher level cognitive skills (Wiggins, 1990). Typically students must demonstrate what they have learned in projects that help them apply principles, produce products, or engage with the material in performances that can be assessed by means other than standardized or objective testing. These situations approximate how scientists do their work (American Association for

the Advancement of Science, 1993; National Research Council, 1996). The Assessment Task Force proposed the concept of authentic teaching as the use of real-life examples and problems to teach the principles of behavior. Authentic teachers provide instruction to help students generalize what they have learned in the classroom to situations they will find in daily life. Selecting real-life examples that are realistically linked to the student experience enhances the likelihood that students will be able to absorb, integrate, and apply practical examples to solving problems scientifically. Furthermore, these practical applications should focus student interest more strongly on the content of psychology. The relevance of scientific thinking should perpetuate greater reliance on using scientific problem solving beyond the termination of the course. The rubric permits generation of performance criteria in authentic assessments that employ authentic teaching applications.

Evaluating Psychology Courses

Good psychology courses flow from well-designed syllabi that articulate expected outcomes (Appleby, 1999). Psychology courses that share in the mission of promoting scientific inquiry can use the rubric to guide course development and subsequently reinforce good teaching and learning practices. This practice may be especially helpful in college programs that have designed capstone courses to integrate undergraduate learning experiences as recommended by the St. Mary's Conference (Brewer et al., 1993). The highest level of development can provide a structure from which to evaluate a senior portfolio; however, course goals and objectives throughout the curriculum can be compared to the rubric's developmental expectations.

Evaluating Psychology Programs

Adoption of this developmental rubric can provide an important evaluation tool to promote program coherence. The rubric facilitates the articulation of objectives for courses and curricula. As departments undergo self-study for program review, faculty can use the rubric as a resource for evaluating how the courses make contributions to overall development as well as how they fit together. Recognition of problems with that development may argue for tighter sequencing of courses or changes to course requirements. More strategic requirements might reduce the common complaints that students have regarding repetition among their courses or the lack of relevance of their learning to their lives.

Implications of Rubric Use

The implications of applying the developmental rubric in authentic teaching and learning contexts are multifaceted. This developmental rubric has the advantage of clearly defining the precise goals of instruction for students, teachers, programs, and external entities interested in the outcomes of psychology education. We offer an analysis of the implica-

tions that addresses the perspectives of students, teachers, departments, and other educational stakeholders.

Why Should Students Care?

Use of the rubric should have personal value for students. Applying course content and reasoning to problems that students experience as real should enhance their interest in learning the content and principles of psychology. Students are likely to see more value in their course work and strive harder to master tasks presented by the teacher by applying their growing knowledge and proficiency to tackle problems they have experienced, issues they have seen in the media, or challenges they may expect to encounter in their future work environments. Thus, use of the rubric in authentic assessment strategies should do more to assure valid student performances than more traditional procedures.

When teachers specify performance criteria in authentic assessment practices, students have greater opportunity to become partners in the process. Teachers can propose, debate, and modify criteria to reflect student input. This partnership facilitates a sense of empowerment with students seeing themselves as sharing responsibility for their success throughout a course or program. These conditions reduce resistance to learning or the work associated with it because objectives are clear and have been shared. Such openness facilitates students feeling more personal control over the probability of academic success.

We recommend sharing performance criteria with students to enable active self-assessment on any given assignment based on the rubric. When students can compare their judgments about what they have achieved with the perspective of the teacher, their self-assessment skills should sharpen over time. As a consequence of incorporating self-assessment more routinely in student evaluation, students are better prepared to talk about what they have accomplished as learners in psychology. This practice will help them advocate more effectively for their professional goals.

This developmental rubric can be a useful formative tool for students, reflecting where they have been, illustrating where they are, and pointing to where they are going. Assessment with the rubric, then, provides students with a practical road map permitting them to evaluate progress and the distance left to travel. However, we caution teachers against sharing the entire rubric with students who are in the early stages of their education because we fear that they will feel overwhelmed. Introducing the complete rubric as students begin to enter more specialized courses in the major, such as experimental design or theory-based courses, may produce the greatest appreciation for the value of the road map.

Why Should Psychology Teachers Care?

Beyond its motivating influence on students, use of the scientific inquiry rubric offers many other practical benefits for teachers. Designing courses and assignments around the rubric transcends teaching the content of psychology. Authentic assessment embeds learning content in a plan to influence the development of sophistication in thinking and

performance. The clarity of goals produced by applying the rubric to exercises should make teaching an easier and more rewarding activity. The developmental framework provided in the rubric guides the expectations a teacher may have for students, helping to select exercises and set criteria for successful performance. The rubric defines progress clearly so that teachers may observe student improvement in behavioral steps and derive the enjoyment of seeing the productive outcome of instruction. Furthermore, the unequivocal identification of behavioral expectations for students enhances their likelihood of success, in turn enhancing the pleasure in teaching and performance evaluation.

Student evaluation should involve more than just grade determination. Effective teaching and assessment practices need to take into account the development of the whole student (Mentkowski et al., 2000). Students can receive comprehensive feedback about their performance when teachers use criteria sheets to evaluate student performance. Teachers can effectively check off successful aspects of performance, leaving time to write a positive note or a specific recommendation for change. Additionally, requiring students to self-assess the quality of their work before submission further diminishes the teacher's feedback burden. If the criteria are sound and easy to apply, the teacher can confirm the student's self-assessment.

Caught up in the minutiae of teaching, teachers can easily lose sight of long-term student development objectives. The rubric encourages teachers to practice their craft metacognitively, leading them to consider more explicitly the relation of teaching exercises to the thought processes they are trying to foster. Targeting the appropriate level of student thinking can guide selection of instructional methods and prevent a teacher from creating unchallenging or unrealistic performance demands. In addition, having the rubric as a curricular backdrop provides an important foundation for faculty members who are launching their teaching careers to understand the larger context to which their classes will contribute.

Reliance on the rubric can also assist with routine classroom aggravations. For example, teachers working with beginning students sometimes feel taxed by their questions that demand simplistic answers or by their passivity in response to a perceived authority (cf. Perry, 1970). Indicators from the rubric suggest that such simplistic approaches are developmentally appropriate. As a consequence, psychology teachers can show greater patience with unsophisticated student responses and begin to plan ways to help the students move to the next level. Furthermore, recognizing student progress in moving beyond acceptance of simplistic explanations can help to generate respect for student achievement and satisfaction as well as the teacher's role in guiding their learning.

Students sometimes question the relevance of what teachers ask them to learn. If assignments reflect expectations set forth by the rubric, teachers should be able to justify the role that a particular assignment or course will play in the overall development of the student's scientific inquiry skills and psychological sophistication.

Teachers can use this scientific inquiry rubric as a diagnostic tool to evaluate what is missing from student performance. By identifying and applying behavioral criteria, teachers can pinpoint areas of deficiency in student perfor-

mance and begin to address those problems. Because assignments relate to specific rubric indicators, the teacher can describe concretely, using terminology from the rubric, how the student will need to change to be successful.

The indicators developed in the rubric can give some important guidance on how to provide meaningful feedback to students about their life decisions. For example, if a student professes a desire to go to graduate school but fails to demonstrate any of the behavioral criteria associated with professional performance, the teacher can use specific language that can soften the blow of a declined letter of reference and perhaps lead to richer discussions of other more appropriate job choices. The performance criteria may also be useful in making other important academic distinctions, such as honors designations or scholarship qualifications.

The rubric can also facilitate the teacher's self-assessment. Use of the rubric may help the teacher determine which instructional methods have been generally successful in advancing the developmental progression of students and which have not. In addition, teachers can compare their behavioral objectives with the rubric's indicators for a course at a particular level to determine if their expectations are reasonable. Recognizing gaps in course designs should help teachers adjust their instructional methods in areas needing modification. Ultimately, these refinements should result in more effective teaching and learning.

Teachers may also benefit from using the developmental rubric when they provide evidence of their teaching effectiveness (Edgerton, Hutchings, & Quinlan, 1991; Seldin, 1997). Aligning authentic teaching methods with behavioral outcomes in the rubric permits an ongoing collection of data showing advancement toward stated goals. A teaching portfolio can combine information about authentic instruction methods and assessment outcomes from the rubric. This information provides a more direct connection between teaching evaluation and student performance. In essence, routine student evaluation with the rubric also becomes the data for external assessment of the teacher, a time-saving arrangement. The portfolio links teaching activities and demonstrates on a continuing basis the teacher's accountability for student learning.

Although the advantages of adopting authentic assessment strategies are numerous, we would be remiss if we did not address some disadvantages as well. Many teachers may find authentic assessment requires more energy and time. They may perceive the approach negatively as just one more thing to learn. In addition, such practices may render the authentic assessment practitioner as an outlier in the department. Unless reinforced by the many potential gains involved in the practice, enthusiasm may fade over time. However, even without full implementation of authentic assessment, we believe that familiarity with the rubric will lead to teacher improvement.

Why Should Psychology Programs Care?

Adoption of this developmental rubric also has important implications for program and curriculum development. Whereas individual courses may apply only one developmental level appropriate to their position in a curriculum, pro-

grams should examine the full range of the rubric to reflect on and promote curricular coherence. Programs at the level of high schools, 2-year and 4-year colleges, and graduate schools may all identify the point in the rubric at which typical students enter and the point to which they should progress before exiting. Discussions about the ideal curriculum can unfold to help chart student activity in the interim. Indeed, cooperative application of the rubric across educational contexts could facilitate articulation of credit and placement as students move from high school to college and from 2- to 4-year colleges.

The rubric offers support for sequential development within the curriculum and encourages analysis of proper placement of particular courses in a curriculum. Thoughtful program evaluation using the progression in the developmental rubric as a model may point the way to a specific sequential redesign of curricula. Creation of a rational course sequence based on realistic understanding of student development permits the opportunity for courses to build on each other, later courses taking advantage not just of the content learned but of the cognitive advancement promoted in earlier courses.

With the rubric as the backbone of a developmental curriculum, additional benefits accrue. Designing an integrated curriculum is very much a cooperative and collaborative endeavor, facilitating the ability of all teachers involved in the program to understand and appreciate what is happening in their colleagues' courses. Working this way toward a set of agreed-on goals has the potential to enhance departmental cohesiveness (Sherif, Harvey, White, Hood, & Sherif, 1961). Ultimately, this approach will yield courses and curricula with better articulated objectives and outcomes. Clear identification of where particular concepts should be introduced in the curriculum will help to eliminate unnecessary redundancy or highlight areas where planned redundancy will reinforce student learning. In short, applying the developmental rubric to curriculum development will promote cohesion by reducing curriculum chaos.

Sharing the conceptual foundations of the department's course requirements with students can facilitate more effective advising. Presentation of the systematic character of the program, its goals, and its projected outcomes to students should help them to see and accept the rationality of the curriculum. Students can also be invited into the process and actively evaluate their progress through the program as well as offer constructive criticism about the way in which curricula meet those objectives.

Conclusions

Outcomes assessment has become a standard feature of doing educational business. Reliance on authentic assessment can shift the emphasis away from high-stakes, standardized testing that primarily examines factual content knowledge in a traditional, sometimes artificial manner toward the development of performance assessment strategies that clearly reflect skills applied to practical problems. For example, collection of these products in an evolving portfolio affords a richer record of achievement than is obtained with

factual testing. Furthermore, reliance on applications of the rubric shifts the specification of instructional goals and outcomes from often-faceless constituents external to the institution toward the teacher and department. Such a shift places responsibility for progress clearly in the realm of the cooperative interaction between teacher and student.

Teachers of psychology at all levels are inherently involved in a partnership to maximize student progress toward sophisticated scientific inquiry in psychology. Adoption of the rubric provides a framework for facilitating dialogues among these teachers. The community of psychology teachers can share specific pedagogical strategies to promote student development. However, their recognition of the interdependence of outcomes and instruction at all levels of psychology education can encourage them to think outside the traditional boundaries of their areas of responsibility. They can more easily participate in local conversations about creating successful, seamless transitions between levels. Working from a common model of cognitive development should foster greater understanding and appreciation among levels, contributing to improved psychology education overall. Teachers of psychology also can highlight the importance of each assignment, each course, and each teacher for contributing to students reaching the level of success to which they aspire. Placing teachers' roles in this larger context provides concrete evidence of the value of individual teachers' contributions to student development and advancing psychology.

The proposal of the developmental rubric in conjunction with authentic teaching and assessment that we present in this article clearly suggests a beginning, not a fully packaged solution. We are hopeful that it will serve as the foundation for further research on the effectiveness of using this developmental rubric and for sharing authentic applications of it from high school through graduate training. Additionally, we recognize that this research may illuminate additional components for the rubric and encourage work on developmental rubrics in other domains that teachers identify as essential to psychology education.

Among the potential areas of research that emerge from the implications we have discussed here include evaluation of student affect toward this style of teaching, examination of teacher satisfaction with outcomes achieved, and determination of levels of mastery of content and procedural knowledge with this approach. Perhaps the most apparent research question is whether knowledge and abilities gained at each stage in the rubric will transfer or generalize beyond the classroom. We invite discussion of the developmental research strategies that will help to determine the value of our proposals.

As opportunity for discussion of our ideas emerges at conferences and in print, we expect to be able to compile a broad set of examples of authentic teaching and assessment that we will make available to psychology teachers. We are eager for discussions of these methods to begin, and we extend our partnership to invite all psychology teachers to become involved in them. To facilitate this discussion and to make more concrete the form that authentic teaching and assessment may take when using the developmental rubric, we offer in an appendix to this article one application of these principles. This example, "The Psychology Newspaper," was originally designed for use in a capstone history of psychology

course. As noted in the appendix, however, the assignment is adaptable to a wide variety of courses and levels and may help a broad cross-section of teachers to begin to work with the developmental rubric.

References

- American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.
- American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.
- Andreoli Mathie, V. A., & Ernst, R. (1999, September–October). The national forum on psychology partnerships: From vision to reality. *PTN: Psychology Teacher Network*, 9(4), pp. 1, 5, 8, 16.
- Appleby, D. (1999). How to improve your teaching with the course syllabus. In B. Perlman, L. I. McCann, & S. H. McGadden (Eds.), *Lessons learned: Practical advice for the teaching of psychology* (pp. 19–24). Washington, DC: American Psychological Society.
- Astin, A. W. (1993). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education*. New York: Macmillan.
- Banta, T. W., Lund, J. P., Black, K. E., & Oblander, F. W. (1996). *Assessment in practice: Putting principles to work on college campuses*. San Francisco: Jossey-Bass.
- Barr, R. B., & Tagg, J. (1995). From teaching to learning: A new paradigm for undergraduate education. *Change*, 27(6), 12–25.
- Brewer, C. L., Hopkins, J. R., Kimble, G. A., Matlin, M. W., McCann, L. I., McNeil, O. V., Nodine, B. F., Quinn, V. N., & Sandra. (1993). Curriculum. In T. V. McGovern (Ed.), *Handbook for enhancing undergraduate education in psychology* (pp. 161–182). Washington, DC: American Psychological Association.
- Bryant, W. H. M., & Benjamin, L. T., Jr. (1999). Read all about it! Wundt opens psychology lab: A newspaper assignment for history of psychology. In L. T. Benjamin, Jr., B. F. Nodine, R. M. Ernst, & C. T. Blair-Broeker (Eds.), *Activities handbook for the teaching of psychology* (Vol. 4, pp. 47–49). Washington, DC: American Psychological Association.
- Dunn, D. S., & McCarthy, M. A. (2001, February). *Assessing student outcomes: A rubric for scientific reasoning*. Paper presented at the annual Southeastern Conference on the Teaching of Psychology, Kennesaw, GA.
- Edgerton, R., Hutchings, P., & Quinlan, K. (1991). *The teaching portfolio: Capturing the scholarship in teaching*. Washington, DC: American Association of Higher Education.
- Halpern, D. F. (1988). Assessing student outcomes for psychology majors. *Teaching of Psychology*, 15, 181–186.
- Hogan, K. (1999). Thinking aloud together: A test of intervention to foster students' collaborative scientific reasoning. *Journal of Research in Science Teaching*, 36, 1085–1109.
- Hogan, K., & Fisherkeller, J. (2000). Dialogue as data: Assessing students' scientific reasoning with interactive protocols. In J. J. Mintzes, J. H. Wandersee, & J. D. Novak (Eds.), *Assessing science understanding: A human constructivist view* (pp. 95–127). San Diego, CA: Academic.
- Kitchener, K. S., & Fischer, K. W. (1990). A skill approach to the development of reflective thinking. In D. Kuhn (Ed.), *Developmental perspectives on teaching and learning* (Vol. 21, pp. 48–62). Basel, Switzerland: Karger.
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Linn, J. E., & Gronlund, M. A. (1995). *Measurement and assessment in teaching*. Englewood Cliffs, NJ: Merrill.

- Maki, P. L. (2001). From standardized tests to alternative methods: Some current resources on methods to assess learning in general education. *Change*, 33(2), 29–31.
- McGovern, T. V., Furumoto, L., Halpern, D. F., Kimble, G. A., & McKeachie, W. J. (1991). Liberal education, study in depth, and the arts and sciences major—Psychology. *American Psychologist*, 46, 598–605.
- Mentkowski, M., Rogers, G., Doherty, A., Loacker, G., Hart, J. R., Richards, W., O'Brien, K., Riordan, T., Sharkey, S., Cromwell, L., Diez, M., Bartels, J., & Roth, J. (2000). *Learning that lasts: Integrating learning, development, and performance in college and beyond*. San Francisco, CA: Jossey-Bass.
- Morgan, B. L., & Johnson, E. J. (1997). Using a senior seminar for assessing the major. *Teaching of Psychology*, 24, 156–159.
- National Research Council. (1996). *National science standards*. Washington, DC: National Academy Press.
- O'Neil, J. (1992). Putting performance assessment to the test. *Educational Leadership*, 49(8), 14–19.
- Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco: Jossey-Bass.
- Perry, W. G., Jr. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. Austin, TX: Holt, Reinhart, & Winston.
- Piaget, J. (1963). *The origins of intelligence in children*. New York: Norton.
- Schneider, C. G., & Schoenberg, R. (1998). *The academy in transition: Contemporary understandings of liberal education*. Washington, DC: American Association of Colleges and Universities.
- Seldin, P. (1997). *The teaching portfolio: A practical guide to improved performance and promotion/tenure decisions* (2nd ed.). Bolton, MA: Anker.
- Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1961). *Intergroup cooperation and competition: The robbers cave experiment*. Norman: University of Oklahoma Press.
- Skinner, B. F. (1965). *Science and human behavior*. New York: Free Press.
- Strong, S., & Sexton, L. C. (1996). Performance assessment for state accountability: Proceed with caution. *Journal of Instructional Psychology*, 23, 68–74.
- Trice, A. (2000). *A handbook of classroom assessment*. New York: Addison-Wesley-Longman.
- Watson, R. I. (1967). Psychology: A prescriptive science. *American Psychologist*, 22, 435–443.
- Wiggins, G. (1990). *The case for authentic assessment: ERIC digest*. (ERIC Document Reproduction Service No. ED 328611)
- Wlodkowski, R. J., & Ginsberg, M. B. (1995). *Diversity and motivation: Culturally responsive teaching*. San Francisco: Jossey-Bass.
- Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review*, 20, 99–149.

Appendix History of Psychology The Psychology Newspaper

Context

History of Psychology often serves as a capstone course in the psychology major. The course should provide an opportunity for students to integrate their skills across the domains of scientific inquiry. Bryant and Benjamin (1999) described the basic design for this assignment; J. Halonen developed the behavioral criteria.

Purpose of Assignment

To collaborate with others in the production of a newspaper that summarizes the important events in psychology along with other “current” events for a selected year. Students sometimes struggle with understanding the significance of historical events in psychology. This activity helps them place the critical events of psychology in a broader cultural and historical context as they refine their collaboration skills.

Directions

Join with a group of 4 or 5 students. Select a particular year in psychology’s history and develop a newspaper page that communicates the primary events of interest in psychology during that year along with other events that happened during the year. Some class time can be used to organize the tasks of the group members, but the majority of the work will need to take place independently in the library. Display final products during a specified class period. Be prepared to discuss the role of culture and history in speculating why the psychology events occurred during your assigned year. Complete a self-assessment on the quality of your work and give feedback to other students based on the grading rubric.

Rubric for Newspaper Quality and Presentation: Assessment of Product

Description skills: Observation.

___ Summarizes key ideas accurately

Conceptualization skills: Concept application.

___ Accurately identifies psychological events with assigned year

Communication skills: Resource-gathering skills, conventional expression.

___ Uses resources selectively and appropriately

___ Balances psychology events with other newsworthy events

___ Expresses events in language appropriate for newspaper audience

___ Produces aesthetically pleasing newspaper

Problem-solving skills: Bias detection and management.

___ Speculates about how cultural and historical influences may have influenced psychology’s developments

Rubric for Collaboration: Assessment of Individual Contribution to Group Process

Collaboration: Project completion skills, process management, leadership, consensus-building skills.

___ Completes complex project with others with minimal direction

___ Contributes equitably to effectiveness of the group through leadership, cooperation, and follow-through on commitments

___ Evaluates quality of group process to improve work

___ Offers developmental feedback to group members on quality of performance

Rubric for Self-Assessment

Self-assessment: Self-regulation, self-reflection.

- ___ Accurately characterizes own contributions to group process
- ___ Evaluates effectiveness of group plan in achieving desired outcome
- ___ Describes lessons learned about working with groups over time that may influence future collaboration

Grading Criteria

Three factors determine individual grades for this project: the quality of newspaper page and its presentation during the class, the quality of the collaborative effort, and the quality of self-assessment. The relative weights of the available points assigned for this project include 50% for the newspaper and class presentation, 30% for collaboration, and 20% for the self-assessment.

The newspaper pages themselves tend to be high-quality, competitive productions. The presentation of the newspaper page gives the group an opportunity to explain and defend the choices made by the group.

The collaborative score can be more challenging to determine because some students sometimes do not perceive their contribution in the same manner as the rest of the group. Collecting impressions from group members on individual performance provides an opportunity to weight collaborative contributions differentially. In addition, students who have been lackluster in their performance by consensus of their peers must deal with the discrepancy in their self-assessments.

Broader Scope of This Assignment

Instructors in many topical areas at different levels may adapt this assignment to their courses. For example, child behavior courses offered at the “Basic” or “Developing” levels may require a newspaper composed of articles from a variety of news areas (e.g., economics, nutrition, health) that have an influence on children’s development. The instructor could then select the behavioral criteria from the rubric that are relevant for the level of the course.

Notes

1. Financial support was provided by the Carnegie Foundation through a CASTL grant.
2. The Assessment All-Stars of the Psychology Partnerships Project (P3) have had help from many partners along the way. We acknowledge Virginia Andreoli Mathie and the P3 Steering Committee for their design that brought the original team together and offered encouragement and support throughout our commitment. We had helpful feedback from several teaching communities: the Southeastern Teaching of Psychology Conference, the University of Green Bay Psychology Teachers’ Conference, the Steering Committee of the Rhode Island Teachers of Psychology, and the Northern Kentucky University High School Psychology Teachers’ Institute. Several individuals gave us feedback on various drafts of the article and the rubric, including Paul C. Smith, Rebecca McKenzie, and Monica Reis-Bergan. We are also grateful to the *Teaching of Psychology* reviewers for their excellent suggestions.
3. Send correspondence to Jane Halonen, School of Psychology, James Madison University, MSC 7401, Harrisonburg, VA 22807; e-mail: halonejx@jmu.edu.