

Measuring beyond content: a rubric bank for assessing skills in authentic research assignments in the sciences†

Tara L. S. Kishbaugh,^{*a} Stephen Cessna,^a S. Jeanne Horst,^b Lori Leaman,^c Toni Flanagan,^c Doug Graber Neufeld^d and Matthew Siderhurst^a

Received 13th December 2011, Accepted 17th February 2012

DOI: 10.1039/c2rp00023g

Herein we report the development of an analytic rubric bank to assess non-content learning, namely higher order cognitive skills, the understanding of the nature of science, and effective scientific communication skills in student research projects. Preliminary findings indicate that use of this tool enhances our students' learning in these areas, particularly in the area of communication, as well as eases and speeds the assessment of these skills. Our rubric bank is notable for its adaptability to a wide range of assignments, developmental levels, and science courses.

Introduction

What should science students learn?

Higher education science educators increasingly recognize that becoming a scientifically literate member of society requires much more than just learning the content of our disciplines (Fink, 2003; McCray *et al.*, 2003; AACU, 2005). In fact numerous science education reform documents encourage the broadening of learning objectives beyond content and towards scientific literacy (AAAS, 1993; NRC, 1996; NRC, 2011).

While varying definitions of scientific literacy could be stated, learning objectives can roughly be classified into three categories: (1) the development of higher order cognitive skills, (2) the formation of an articulate understanding of the nature of science and scientific inquiry, and (3) the establishment of skills in oral and written communication of scientific information (AAAS, 1993; NRC, 1996; McCray *et al.*, 2003; ACS, 2008; AAAS, 2009).

The first of these broad categories, higher order cognitive skills (abbreviated 'HOCS'), encompasses ideas of scientific problem solving skills and critical thinking skills (Zoller, 1993). While various definitions of HOCS and critical thinking skills have been

posited and often debated (Ennis, 1985), the notion generally stems from Bloom's taxonomy of educational objectives (1956), and many useful definitions for teaching HOCS still use Bloom's original ideas (AAAS, 1993; Bissell and Lemons, 2006). Bloom noted six learning objectives, beginning with two 'lower order' objectives: the knowledge of and the understanding of factual content matter. Four other learning objectives are considered 'higher order'; they build on knowledge and understanding of the content, but also require critical thinking skills. Bloom's four higher order cognitive skills are:

- *application* (generalizing),
- *analysis* (breaking problems down to their component parts),
- *synthesis* (making connections between the parts of different problems), and
- *evaluation* (assessing information and making value and truth judgments).

HOCS are the critical components of scientific problem solving and are the basis from which scientific truth claims are made and evaluated (AAAS, 2009). Thus, the benefits of students learning HOCS include not only their success in later science coursework and eventually greater employment opportunities, but also greater scientific literacy for informed participation in a democratic society (Hand *et al.*, 1999; Quitadamo *et al.*, 2008).

Although research indicates that gains in critical thinking are achieved when faculty provide purposeful instruction in critical thinking (Halpern, 1996; Pascarella and Terenzini, 2005), this opportunity may be underutilized in science departments (Jacobs, 2004). Traditional science education has been one of "received knowledge" through lecture-oriented teaching and the collection of facts, failing to capitalize on opportunities for the development of HOCS in college students (Zoller, 1999). The challenge remains for science educators to develop critical thinking skills in college students in such a way that students are able to transfer those skills to

^a Chemistry Department, Eastern Mennonite University, 1200 Park Road, Harrisonburg, Va, USA. E-mail: tara.kishbaugh@emu.edu, cessnas@emu.edu, matthew.siderhurst@emu.edu; Fax: 01 540 432 4488; Tel: 01 540 432 4465

^b Psychology Department, Eastern Mennonite University, 1200 Park Road, Harrisonburg, Va, USA. E-mail: jeanne.horst@emu.edu; Fax: 01 540 432 4488; Tel: 01 540 432 4431

^c Education Department, Eastern Mennonite University, 1200 Park Road, Harrisonburg, Va, USA. E-mail: lori.leaman@emu.edu, shaneton@wispwest.net; Fax: 01 540 432 4071; Tel: 01 540 435 4145

^d Biology Department, Eastern Mennonite University, 1200 Park Road, Harrisonburg, Va, USA. E-mail: neufeld@emu.edu; Fax: 01 540 432 4488; Tel: 01 540 432 4401

† Electronic supplementary information (ESI) available: Full rubric bank, examples of assignment specific rubrics and full survey results. See DOI: 10.1039/c2rp00023g

novel learning situations, across courses and within the context of real-world, ill-structured problems (Halpern, 2001).

The second broad category of science literacy, the nature of science (NOS), refers to the epistemological underpinnings of the scientific endeavour (McComas, 1998; Lederman, 2007). NOS concepts include answers to such questions as: How do scientists gain knowledge? What makes scientific knowledge valid? Does scientific knowledge change? What are the methods of scientific inquiry? What is the relationship of science to society? Answers to these questions are embedded in the philosophy and sociology of science, and so vary from 'modern' or 'postmodern' positions (these philosophical issues are outside of the focus of this paper, but see Good and Shymansky (2001), Wong and Hodson (2009), and Lederman (2007)). While one coherent vision of NOS is thus unlikely to be widely accepted, seven non-controversial facets of NOS have been tested against numerous philosophers of science and practicing PhD research scientists (Wong and Hodson, 2009), and have been widely adopted in the science education literature (Lederman, 1992, 2007).

The nature of science is

- *empirical* (scientific knowledge is based on and/or derived from observations of the natural world),
- *tentative* (scientific knowledge is subject to change with new observations and with the reinterpretations of existing observations),
- *inferential* (scientific knowledge is based on observations, and also on inferences drawn from those observations; there is a critical distinction between scientific claims and the evidence on which such claims are based),
- *theory-laden* (scientific knowledge and investigation are influenced by scientists' prior theoretical and disciplinary commitments),
- *embedded in a wider culture* (science is a human enterprise, practiced within and affecting society and culture; scientists are influenced by culture—in their beliefs, values, norms, and prior knowledge),
- *founded on no specific scientific method* (there is no universal step-wise method that guarantees the generation of valid scientific knowledge; many different methodologies are valid means of scientific knowledge formation and contribute together to validate theories), and
- *creative* (science is a creative process; scientific concepts, such as atoms or species, are creative and useful models, not perfect copies of reality).

Students enter our science classrooms and teaching laboratories with numerous misconceptions regarding these NOS concepts (Marchlewicz and Wink, 2011). For example, a misconception of 'the scientific method' is strongly reinforced in many chemistry textbooks (Abd-El-Khalick *et al.*, 2008); frequently the first chapter of high school and introductory college science textbooks will present 'The Scientific Method' as a rigid step-wise procedure, starting with observations, moving to hypothesis, and leading inevitably to experimentation, which will either confirm or reject the hypothesis. However, in practice, science moves forward on the basis of multiple methods, many of which are not experimental, nor hypothesis-driven (*e.g.* chemical synthesis, comparative genomics, and theoretical physics (Wong and Hodson, 2009)). The major science education reform documents urge the teaching of NOS concepts as a critical part of scientific

literacy development, in both science majors' and non-science majors' coursework (AAAS, 1993; NRC, 1996).

In addition to learning HOCS and NOS concepts, the third broad category of literacy that college science students should learn are skills for communicating science. The benefits of science students gaining communication skills specific to our disciplines (keeping a laboratory notebook, writing research papers, giving oral presentations, and presenting posters) are self-evident, for the benefit of industry, medicine, education and research, in which our students will eventually participate (AAAS, 1993; ACS, 2008). Communicating effectively within the accepted traditions of the scientific culture and enterprise also serves to reinforce skills of argumentation and evaluation (HOCS), and NOS conceptions (Jiménez-Aleixandre *et al.*, 2000). Thus, while a categorization of learning objectives into these three broad groups might be useful, it should be noted that significant overlap between these categories of learning objectives exists, and successful scientists demonstrate HOCS and NOS understanding in their writing and speaking.

Undergraduate science faculty see the benefits of teaching towards these non-content learning objectives, but also note that they are much more difficult to teach and assess than content learning (Bissell and Lemons, 2006). Thus, science faculty could benefit from materials that improve student learning in these areas and provide an easy way to assess if science students meet these non-content course objectives.

Authentic scientific practice

One teaching practice that has been promoted for facilitating learning beyond the content is authentic research-based laboratory assignments (Quitadamo *et al.*, 2008; Russell and Weaver, 2011). More than twenty years ago, Brown, Collins and Duguid posited a new theoretical model of learning, which they termed 'situated cognition' (1989). They surmised that better learning, or at least more practical learning (*i.e.* learning that is more readily transferred to the workplace, community, and household), occurs when the content is situated in a rich real-world context. Situated learning theory applies an ancient apprenticeship model of learning where students learn in an 'authentic community of practice', and is supported by a robust modern cognitive theory (Brown *et al.*, 1989; Lave and Wenger, 1991). In curricula informed by situated learning theory "[t]he individual learner is not gaining a discrete body of abstract knowledge which (s)he will then transport and reapply in later contexts. Instead, (s)he acquires the skill to perform by actually engaging in the process..." (Lave and Wenger, 1991). Thus, learning practical scientific skills (*i.e.* HOCS) requires the *doing* of science, in a manner that moves beyond lectures and confirmatory laboratory projects, and into student engagement in the doing of authentic science research (Singer *et al.*, 2000).

Undergraduate involvement in authentic research has thus been encouraged as a means of improving just the type of beyond-the-content learning that we hope for (NRC, 1996; ACS, 2008). However, providing authentic research apprenticeships for all STEM majors and non-majors (particularly preK-12 preservice teachers) *via* placement with faculty research mentors is unrealistic given the overwhelming numbers. Furthermore, direct evidence of

learning gains in HOCS, NOS understanding and communication skills from unstructured research experiences is lacking (Kardash, 2000; Schwartz *et al.*, 2004; Hunter *et al.*, 2007; Quitadamo *et al.*, 2008). In fact, NOS learning through engagement in authentic research has been demonstrated to be most successful when coupled with explicit instruction and reflection on NOS concepts (Schwartz *et al.*, 2004; Lederman, 2007). The simple introduction of research projects to the undergraduate curriculum is likely not enough to generate substantial learning gains. Thus, while increasing the number and variety of authentic problem-based research projects in several of our chemistry and biology courses, we sought means of facilitating concise and explicit instructor feedback and student reflection with regard to those assignments. We therefore developed an extensive rubric bank as a means of assessing for learning beyond the content.

Formative/summative assessment and rubrics

In the current study, we focused on rubrics as tools that could help us meet several forms of assessment (H. G. Andrade, 2000; Earl, 2003; Andrade and Du, 2005; H. G. Andrade, 2005). Applying rubrics for classroom use shapes teaching beyond merely assigning grades (*i.e.*, summative assessment) by encouraging (1) evaluation of teaching goals, otherwise known as formative assessment (Black and Wiliam, 1998; Carless, 2007), (2) adjustment of assessment for students' developmental level, and (3) tailoring of instruction toward meeting students' needs. More importantly, from the "assessment as learning" perspective, the rubrics provide students with a means for assessing their own work by comparing and adjusting their work against a standard (Black and Wiliam, 1998; H. G. Andrade, 2000; Earl, 2003).

Despite the fact that rubrics have become regarded as effective teaching tools (Reddy and Andrade, 2009), they vary greatly in their utility across assignments and disciplines, and in their ability to impact student learning (Popham, 1997). Moreover, there are few published rubrics that are specific to the science disciplines in higher education, and those available vary greatly in their design and utility. Some are specific for skills (Fitch, 2007), or assignments (Ruiz Primo *et al.*, 2004; Allen and Tanner, 2006; Burke *et al.*, 2006), or courses (Tariq *et al.*, 1998; Halonen *et al.*, 2003). In other cases, rubrics were developed with specific teaching strategies in mind, such as peer-group evaluations (Hafner and Hafner, 2003) or concept maps (Moni *et al.*, 2005). Although useful within their intended contexts, these rubrics tend to have one or more of the following characteristics that limit their usefulness in a broader sense. First, rubrics designed for specific assignments are difficult to transfer to other types of assignments. This both increases the workload for faculty (developing many rubrics for assignments), and hinders the ability of students to understand the broader goals of skill mastery and transferability of knowledge. Second, rubrics that target general skills do not teach what is unique to understanding science. Rubrics that assess general writing skills, for instance, do not necessarily improve students' understanding of NOS or HOCS. Third, rubrics often give only cursory guidance on what the different achievement levels mean, as opposed to being "full rubrics"

(Allen and Tanner, 2006) with clear criteria on performance. This creates inconsistency as different evaluators interpret the rubric, and also limits the ability of students to clearly understand what the evaluations mean.

Thus, science teachers in higher education who are seeking a tool that is appropriate for their teaching goals have a paucity of options in the published literature, and in particular have few options for rubrics that are designed towards student learning of broader goals of scientific literacy in a variety of assignments. Timmerman *et al.* (2010) published a "universal" rubric aimed at scientific understanding across various disciplines and subjects; it is notable for being a comprehensive rubric that avoids many of the limitations of those that are listed above. However, rubric items are specifically tailored to the traditional sections of a scientific report (*i.e.* introduction, methods, *etc.*) and do not reflect the pervasive nature of HOCS and NOS throughout many types of assignments. Our rubric bank is intended to provide a universal and consistent assessment tool with broad applicability—it can be easily applied both to a variety of assignments, and to a range of disciplines—that targets the broader scientific understandings that are currently emphasized in scientific pedagogy. In addition, this rubric is designed for (and has been used with) both lower and upper level courses, and is thus appropriate for the different developmental stages of learning. Finally, this rubric is also easily used for (and has been used for) peer evaluation, a pedagogical technique with significant learning potential (Hafner and Hafner, 2003; Reddy and Andrade, 2009).

Methodology

Rubric development

We developed nine authentic laboratory research projects across several undergraduate chemistry and biology major and non-major courses in order to enhance the non-content learning of our students. Faculty from the sciences, teacher education and psychology disciplines collaborated to share expertise in the fields of science and assessment/learning in order to develop and evaluate a rubric bank to assess these authentic research projects. The rubric bank was developed to accomplish four over-arching goals:

1. Assess student learning in HOCS, NOS, and Communication skills.
2. Provide consistent assessment of these wide-ranging skills across key chemistry and biology courses taught by various professors throughout the students' four-year program.
3. Allow professors access to a common bank of subskill assessment items applicable to a wide variety of assignments and projects.
4. Use rubrics for summative assessment, formative assessment, and peer review.

In designing the rubric bank, we chose to use the analytic rubric approach (Mertler, 2001), in which instructors clearly articulate levels of performance for specific criteria and then sum the individual scores to obtain a total score. Although development of analytic rubrics can be time consuming, they offer students and instructors substantial feedback (Mertler, 2001), as well as provide detailed results about patterns of

Table 1 Timeline of rubric development

Date	Who	Task	Outcome
May 2009	Assessment faculty	Align research project learning goals with assignments Examine science course syllabi for course learning outcomes	Create skeleton of rubric bank: <i>HOCS</i> , <i>NOS</i> , & <i>Communication skills</i> Choose skills to leave out: specific content, writing, laboratory skills
June 2009	Assessment faculty	Examine key sources for science criteria ^a	Create list of sub-skills
July 2009	Assessment & science faculty	Edit rudimentary rubric	Align with assignments, Find universal language
August 2009	Science faculty & students	Pilot rubric & garner feedback with summer research students	Minor revisions to language Increased awareness of the importance of item choice in creating a usable rubric
Fall 2009	Science faculty	Pilot rubric bank in several science courses	Minor revisions to language
April 2010	Assessment & science faculty	Pilot rubric at a cross disciplinary poster session	See the need for inter-rater reliability study
May 2010	Assessment & science faculty	Inter-rater reliability study	Minor revisions to language Better training/instruction on how to use rubric and how rubric will be used
Fall 2010–Spring 2011	Science faculty	Implement rubric in numerous courses	Determine departmental guidelines for developmentally appropriate work

^a *Standards of curriculum and learning outcomes* from professional associations (ACS, 2010) and various universities, *skill taxonomies* (Bloom, 1956; Anderson *et al.*, 2001), existing *science rubrics*, the *research base* on *HOCS* (Zoller, 2001) and other critical thinking literature.

student achievement and development (Maki, 2004a, 2004b). We modified Maki's process for developing scoring rubrics (2004a, 2004b) as demonstrated in Table 1. Throughout the stages of development, assessment faculty (authors Leaman, Horst and Flanagan) and science faculty (authors Kishbaugh, Cessna, Graber Neufeld and Siderhurst) became more aware of each others' discipline specific language with their accompanying nuances of meaning and application. We met as a team to modify the evolving rubric bank criteria based on feedback from all science faculty members' use of the rubric bank in their courses, to be certain that we were accurately representing the criteria that we wanted to assess and to learn how to use the rubrics in ways that were accurate, fair and consistent (American Educational Research Association, 1999; Andrade, 2005).

During the first phase of the development of the rubric bank, the assessment faculty created a chart that aligned the research project learning goals, laboratory project assignments, and course learning outcomes to create the skeleton of the major sections of the rubric bank:

- Content Knowledge,
- *HOCS*, often used in lab reports, research papers, and the content of posters,
- *Communication Skills* in oral presentations, poster presentations and the written aspect of research papers, and
- *NOS*, rarely measured except on specific assessments intended to measure NOS such as the SUSSI (Student Understandings of Science and Scientific Inquiry) (Liang *et al.*, 2006).

These were expanded to include a subset of skills, listed as individual rows in that section and including a description of the sub-skill for each developmental score, also known as an anchored description (see the full rubric in the supplemental materials). For example, selected rows from the *HOCS* section of the bank include: *Analysis—Clarity of Research Question*, *Analysis—Identifies Rationale, Hypothesis, or Systematic Approach*, *Application—Safety and Ethical Consideration*, *Application—Data Presentation*. The top row of the rubric bank includes a scoring grid ranging from a high developmental

scientific skill level (rating of a 5) to an entry developmental scientific skill level (rating of a 1). We also include a column indicating the recommended scoring cap for developmental level, giving professors the flexibility to tailor the rubric bank for entry-, mid-, and upper-level courses. Another column was added to suggest assignments that commonly include the specific sub-skill. The assessment of *NOS* understanding proved most difficult, so we added specific examples underneath the anchored description to aid us. For example, an assignment demonstrating an entry developmental level understanding of the tentative *NOS* would '*express the naïve view that scientific knowledge is not subject to change*' by:

- 'Failing to acknowledge gaps or misunderstandings, varied interpretations, or controversies;
- Stating the purpose of research as confirming rigid theory or reproducing prior findings; or
- Using naïve language by stating non-tentative conclusions: *e.g.* "prove."

Results

Rubric implementation

One advantage of the rubric bank over other rubrics designed for specific assignments (Ruiz Primo *et al.*, 2004; Allen and Tanner, 2006; Burke *et al.*, 2006) or courses (Tariq *et al.*, 1998; Halonen *et al.*, 2003) is that minimal work is needed to add a rubric in a course; individual rubrics are created by simply selecting appropriate rows from the bank (see the supplemental information for specific poster, writing and oral presentation examples). Because of the size of the rubric bank, it is possible to create a very long rubric; however a rubric that is longer than one page becomes too bulky to be useful. It is also possible to simply choose one item, for example to assess the tentative nature of science on an exam question, or to create a very short three item rubric, such as for oral presentations. Table 2 lists the variety of courses for which a rubric was developed. While the rubric bank was designed to assess non-content learning in

Table 2 Courses and assignments using the rubric bank

Course	Total students	Semesters	Student audience	Assignment styles
Matter and Energy (Introductory chemistry)	100	Fall 2010–2011	Non-science majors (including preK–12 pre-service teachers)	Written research paper
General Chemistry II	50	Spring 2010–2011	First year STEM majors	Research poster; written research paper; peer review
Organic Chemistry II	28	Spring 2010–2011	Sophomore–Junior STEM majors	Research poster; notebooks
Analytical Chemistry	5	Fall 2010	Sophomore–Senior STEM majors	Written research paper
Topics in Chemistry	3	Spring 2011	Junior–Senior STEM majors	Journal reviews
Environmental Toxicology	8	Spring 2010	Sophomore–Junior STEM majors	Research poster
Biochemistry/Chemistry Research and Seminar	20	Fall 2009–2010	Junior–Senior STEM majors	Oral presentation; research poster; peer review
Medicinal Chemistry	10	Spring 2010	Sophomore–Senior STEM majors	Oral presentations; primary literature review; peer review
Foundational Biochemistry	20	Fall 2010–2011	Junior–Senior STEM majors	Written research paper; primary literature presentation
Plant Physiology	8	Spring 2011	Sophomore–Senior STEM majors	Research poster; written research paper
Sustainable agriculture	32	Fall 2009, 2011	Sophomore–Junior STEM majors	Oral presentation; written research proposal

authentic research projects, rubrics have been used for other assignments, such as laboratory notebooks and primary literature reviews.

An example of how we applied the rubric bank for a specific assignment was the rubric created for an annual end-of-the-year cross-disciplinary student research poster session (see supplemental information for the rubric). This rubric is also representative of the iterative changes that the rubric bank has gone through. During the pilot phase, faculty from each discipline (biology, chemistry, biochemistry, and psychology) interpreted and applied the rubric criteria to a group of STEM posters in order to choose the top two. While the raters did not each evaluate every poster, it was clear from the accumulated scores that ratings for certain portions of the rubric bank differed, depending on rater discipline. For example, the practice of including a well-articulated research hypothesis differed among chemistry and biology faculty; indeed sometimes there is no hypothesis (*e.g.* organic synthesis), but there is a clear and organized means of addressing a research problem. When we considered the poster session data in terms of agreement (*i.e.*, “How often did the raters assign the same scores?”) and consistency (*i.e.*, “Did raters apply the scoring criteria consistently across posters?”), we found mixed results across the initial draft of rubric criteria. In order to modify the pilot rubric bank language for greater consistency (American Educational Research Association, 1999) and adaptability across disciplines, we conducted a session during which faculty members blindly and independently rated individual posters from three different courses (general chemistry, organic chemistry, and environmental toxicology) and then discussed rationale for each rating, paying close attention to inconsistent ratings. We continued this process, clarifying the language of the rubric bank, until we had reached adjacent agreement on each of the criteria. In particular the wording for the HOCS analysis sub-skill *Identifies Rationale/Hypothesis* is representative of changes that resulted from this stage of the process.

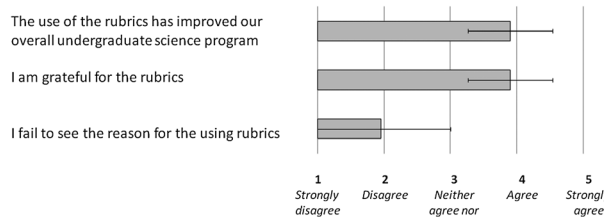
A rubric designed for assessment of research papers in an introductory (non-majors) chemistry course demonstrates an alternate design. This rubric contains some overlap with the one for posters, but has been organized differently to indicate where competencies or sub-skills will be demonstrated within

an assignment, making both communication of the nature of the assignment easier on the students, and summative assessment of those assignments easier on the graders. Some categories are found in specific sections of a poster or a written report, such as the HOCS item *Synthesis* should be found primarily in the conclusions, while other categories, such as the *Empirical Nature of Science*, should be found throughout the work. However, the pervasive and interconnected nature of these skills can also be seen in these rubrics; for example in both the poster and research paper rubric there is a category that describes a HOCS, *the evaluation of the relevance of the study*, as well as a NOS item, *the social and cultural nature of science*. While we created the rubric to assess non-content learning, we included a content knowledge item and referenced our university’s writing rubric rather than creating a new list of sub-skills aimed at grammar and appropriate citations. This reinforces for students the connectivity between the items they view as part of the class (content knowledge) and other areas of scientific literacy. Students in upper level courses (Medicinal Chemistry and Biochemistry/Chemistry Seminar) rated each other’s work using the rubric in a manner that was consistent with the instructors’ scores, supporting the use of the rubric in a peer review process.

Feedback

Use of the rubric bank has had considerable impacts on our teaching and on student learning. To begin to identify and quantify those impacts, we developed two surveys, one for participating science faculty (the biology and chemistry faculty authors of this manuscript), and one for participating students. The presentation of data from this faculty survey is admittedly non-traditional, in that all of the survey respondents are also authors and self-selected; such practice is self-study, a methodology to assist educators in becoming more responsive to the needs of learners (Loughran, 2007; Bullough and Pinnegar, 2007). Thus, we present the data from the faculty survey simply as our collective impression of the rubrics and their use in our courses. In contrast, the student survey data is more in keeping with standard education research practice and so should prove

Student perceptions of the rubrics



Participating faculty perceptions of the rubrics

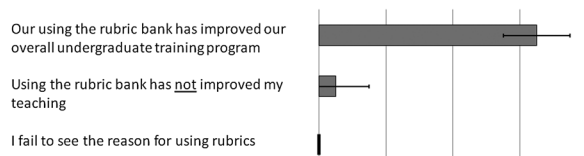


Fig. 1 Student and mean participating science faculty perceptions of the rubrics. The bars summarize responses to the written Likert scale items on the left; participants marked each item with a score from 1 ('strongly disagree') to 5 ('strongly agree'), as indicated. Data shown are means $\pm$ standard deviations of 20 students and four faculty respondents. (All four faculty members marked 'strongly disagree' on the last item.)

useful in terms of drawing generalizable conclusions. Both surveys include Likert-scale and free response items that permitted evaluation of the usefulness of the rubric bank. The surveys were given to the four participating faculty members (we make up 50% of our chemistry and biology departments), and forty-four sophomore-senior level students (80% of the biology or chemistry sophomore-senior majors). Participating students completed at least one course that used the rubric. Paper copies of the student surveys along with a copy of the rubric bank were handed out in several upper-level chemistry and biology courses in the middle of the fall 2011 semester. Student responses were collected confidentially (in keeping with Institutional Review Board guidelines).

Although the student response rate was low (20 students, 45%), we present the results of the student surveys, our collective impressions (via the results of the participating faculty member surveys), and data regarding rubric adoption in our courses to highlight three prominent findings (the entire results of our surveys are available with the supplemental materials).

I. Participating science faculty members and our students view the use of the rubrics as a positive enhancement to our teaching. Of the eight faculty chemistry and biology professors at our university, four (authors Cessna, Kishbaugh, Neufeld, and Siderhurst) were involved in the rubric bank design and implementation. While the rubric bank was originally developed to assess only the products of authentic research projects, we (the participating science faculty members) have been surprised by its usefulness, and have begun using several of the rubric items for other assignments (e.g. laboratory notebooks and primary literature summaries and presentations). Between 2009 and 2011, sixteen unique rubrics were developed from the bank and introduced into all sections of eleven courses, six of which have run multiple semesters, for use on varied assignments, as outlined in Table 2. These eleven courses represent 27% of the unique courses offered in the chemistry (13) and biology (28) departments. The primary reasons why we used the rubric bank so extensively and beyond our original intentions are its ease of use and versatility, namely that it works across several assignment types,

developmental levels of students, and different science courses. We've also found that the process of rubric development has increased our confidence in using the rubric as well as helped us clarify expectations for our students. Thus far, while the rubric has been under development, it has only been used by the four participating faculty. However, after presenting our initial findings to our department colleagues, two other biology faculty members now plan on using items from the rubric bank in their courses in the coming year; we hope that it will be used beyond our departments and beyond our university in the coming years.

Survey results (Fig. 1) further quantify our general enthusiasm for the rubrics, and that of our students. Both the faculty and student survey responses show a perception that the use of the rubrics lead to improvement of the overall undergraduate STEM training program; both surveys also demonstrate a general positive feeling regarding the use of the rubrics.

II. Use of the rubrics eases and speeds grading of complex assignments, while clarifying course objectives; however, use of rubrics also results in the perception of loss of specific useful grading feedback to the students. As noted in Fig. 2, all of the polled students agreed or strongly agreed that the use of the rubrics lead to greater clarity of the learning objectives. Students were not in strong agreement regarding whether grading was faster when performed with the rubrics than with other methods (Fig. 2); however, students were comparing the rates of our grading to those of our colleagues who did not use the rubric—clearly a better controlled study than this is required to draw conclusions regarding student perceptions of grading speed with and without rubrics. All of us (participating faculty) who used the rubrics found that our learning objectives were better clarified for the students, and that grading was made easier and faster (Fig. 2).

Our students also indicate that the feedback provided through grading with the rubrics was not necessarily 'better' than grading without. Fig. 2 shows only weak agreement among students regarding 'better feedback' from assignments graded with the rubrics. While we are relieved that the students' perception of the feedback from grading with rubrics is not apparently 'worse' than grading without them, we had hoped better, in no small part because the learning facilitated through frequent formative assessment is potentially dramatic (see, for example, Andrade, 2000). The use of the ambiguous word 'better' in this survey item is admittedly problematic—what does 'better feedback' mean to the students? To begin to answer this question, students were also given open-ended questions regarding their overall impressions of the rubrics, and the rubrics' facilitation of, or obstruction to the students' perceived learning (complete survey data is included with the supplemental material). Student responses to these questions provide some idea regarding the quality of the feedback from assignments graded with rubrics. There were two types of student responses that appear to speak directly to this question of feedback quality.

First, some students perceive a relative lack of specific comments and instruction when assignments were graded with a rubric, relative to those that were not. For example, one student wrote "[...] upon grading, rubrics tend to leave unanswered questions. I often find myself wondering why I received a certain score in one area due to a lack of explanation."

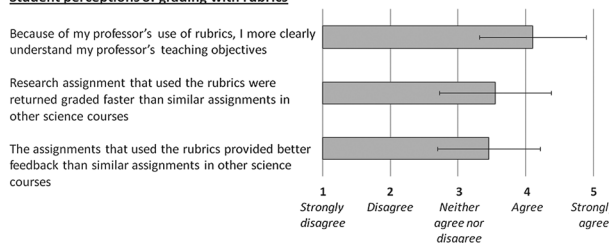
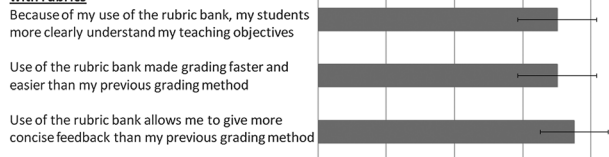
Student perceptions of grading with rubrics**Participating faculty perceptions of grading with rubrics**

Fig. 2 Student and participating faculty member perceptions of grading with the rubrics.

We, the participating faculty, also found this a drawback to rubric use. While we all agree that the rubrics provide more concise feedback (Fig. 2), we also agree that the faster grading provided by rubric use doesn't translate to more valuable feedback. With rubric use, some of us feel like we now tend to give fewer personalized comments, and less reflection on each student's assignment. We hope that this implies an overall beneficial trade-off: if the alternative is more 'cookbook' assignments, then perhaps faster grading of authentic research assignments with some loss of reflection and personalized grading is worth the cost, particularly if rubric use comes with significant other positives such as improved communication of learning objectives and improved quality of student research products (see below).

A second point regarding the quality of the feedback provided by rubric use has to do with the length and complexity of some of the specific rubrics that we designed: students commented that "the rubrics are so long and detailed that I often just glance through them anyway"; "parts of the rubric are complicated and don't always relate to the project". These comments imply that careful rubric design from the rubric bank is critical for students' perceptions of useful feedback, including considerations of length, criteria selection and editing. Perhaps equally important is talking through the rubric with the students prior to the assignment's due date, explaining the criteria and why they were selected for the particular assignment.

III. *Use of the rubric bank has improved the quality of student research products (papers, posters, and oral presentations), but students perceive only minimal changes in learning.* The students generally agree that use of the rubrics has improved their learning, rather than being a hindrance to learning (Fig. 3). Specifically, they perceive an improvement in the quality of their scientific research products (Fig. 3). However, the students are less sure that the use of the rubrics has led to an improvement in their scientific communication skills (Fig. 3). Similarly, when further queried via Likert scale items specifically whether the use of the rubrics lead to enhanced learning of scientific concepts, nature of science understanding, or higher order cognitive skills, the results were inconclusive: the mode answers were '3—neither agree nor disagree', and the standard deviations were large. While we attached the rubric

bank to the survey to refresh the students' memory of what we mean by NOS and HOCS, we are uncertain as to whether the students used this resource or fully understood the wording of these survey items. Thus, we do not present them here in graphical form (see the supplemental information). Taken together, the Likert scale findings suggest that students are generally positive about the learning potential from use of these rubrics, but mostly appreciate them from the perspective that they clarify expectations and allow them to achieve good marks on a given assignment (Fig. 2 and 3).

Consistent with these findings were student responses to open-ended questions; several responses suggest that students perceive that they write better scientific papers in response to the rubrics, but do not necessarily learn more. For example: "Rubrics help because they clearly outline expectations and make it easy to get a good grade."; "[Rubric use] hasn't really changed my learning. [...] I just focus on what the prof wanted." For these student (and several others who gave similar responses), rubrics are not perceived as enhancing learning, but only clarifying expectations for the assignment in order to more readily achieve good marks. This finding is in keeping with those described by Andrade and Du (2005); they found that students appreciate rubrics because they clarify the expectations for a given assignment, but that students may view rubrics solely as a tool for giving teachers what we want, and not as a tool for learning (Andrade and Du, 2005). However, some of our student comments suggest a finer appreciation for the use of the rubrics. For example: "Rubrics are a great help in knowing what exactly is supposed to be done for an assignment and knowing that I'm fulfilling not only class requirements but scientific community standards as well." This underscores the need for professors to reiterate how non-content learning objectives are important to scientific literacy as well as explaining the rubric criteria as part of the assignment.

Our collective thoughts, as the participating faculty members, regarding three years of using of these rubrics and their impacts on student learning are somewhat mixed. We are more confident than our students that their use has led not only to students producing better research products, but also generally to an improvement in their scientific communication skills (Fig. 3).

Use of the rubrics has:

Improved my learning

Hindered my learning

Use of the rubrics has lead to:

Improved quality of student research products (e.g. better research papers and posters)

Improvement in scientific communication skills

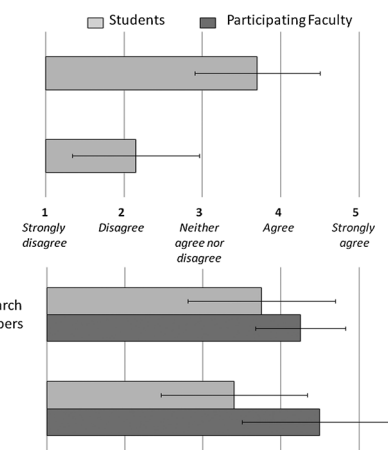


Fig. 3 Student and participating faculty member perceptions of learning gains from using the rubrics.

We are less sure at this point whether the use of the rubrics has lead to appreciable learning of scientific concepts, HOCS, or NOS understanding. We plan to edit the survey to test these items again in the coming semester.

Our hunch is that student perceptions of their learning *via* the use of the rubrics do not necessarily match their actual learning from them (see also Jonsson and Svingby, 2007). Further studies on the effects on learning through student research projects graded with our rubrics are underway.

Conclusions

Teaching beyond the content and towards a scientific literacy that includes higher order thinking skills, nature of science understanding, and skills in scientific communication is an important goal for all of K-16 science education (AAAS, 2009; NRC, 2011). However, teaching towards these broader goals is a challenge.

We have therefore developed a comprehensive and flexible analytical assessment tool for grading undergraduate authentic research products and other assignments, which emphasizes learning beyond content. Our preliminary findings suggest that the use of this tool greatly facilitates students' comprehension of our non-content course objectives. Specifically, both the students and participating faculty members perceive an improvement in scientific writing quality in response to the use of the rubrics. Also, use of the rubrics greatly eases our grading (summative assessment) of those objectives in open-ended assignments, particularly because a rubric bank of common subskills provides greater flexibility in implementation than Timmerman's universal rubric (2010). Moreover, the use of rubrics enabled a more consistent assessment of these wide-ranging skills across chemistry and biology courses taught by various professors throughout the students' four-year program. The process of rubric development informed our instructional practices (formative assessment), and the rubrics promoted assessment as a means of learning, particularly in courses where there were used for peer review. In future work, we hope to establish that real learning gains in NOS understandings, HOCS and scientific communications skills are obtained through research assignments and assessment with these rubrics.

While this project is still a work in progress, we are excited to present our work of the past three years, and hope that the chemistry and biology education community can find our rubric bank useful for several kinds of assignments, potentially from K-16. Some questions still remain, such as how well the rubric would transfer to much larger settings as well as whether such a rubric eases grading by clarifying learning objectives or even facilitates increased use of authentic research projects in the undergraduate setting. We have posted the rubric bank to our webpage, and include it in our supplemental online materials, in hope that it will receive much use.

Acknowledgements

This work was funded by the United States National Science Foundation CCLI grant #0837578 and NSF STEM-STEP grant #0756838. We gratefully acknowledge the participation and feedback of the students in biology and chemistry courses at EMU.

Notes and references

- AAAS, (1993), *Benchmarks for Science Literacy*. New York: Oxford University Press.
- AAAS, (2009), *Benchmarks for Science Literacy*. New York: Oxford University Press.
- AACU, (2005), Liberal Education Outcomes: A Preliminary Report on Student Achievement in College. From http://www.aacu.org/advocacy/pdfs/leap_report_final.pdf.
- Abd-El-Khalick F., Waters M. and Le A. P., (2008), Representations of nature of science in high school chemistry textbooks over the past four decades, *Journal of Research in Science Teaching*, **45**, 835–855.
- ACS, (2008), Undergraduate professional education in chemistry: ACS guidelines and evaluation procedures for bachelor's degree programs. From http://portal.acs.org/portal/PublicWebSite/about/governance/committees/training/acsapproved/degreeprogram/WPCP_008491.
- ACS, (2010), Rigorous undergraduate chemistry programs. From http://portal.acs.org/portal/fileFetch/C/CNBP_024263/pdf/CNBP_024263.pdf.
- Allen D. and Tanner K., (2006), Rubrics: Tools for making learning goals and evaluation criteria explicit for both teachers and learners, *CBE-Life Sciences Education*, **5**, 197–203.
- American Educational Research Association, A. P. A., & National Council on Measurement in Education, (1999), *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Anderson L. W., Krathwohl D. R. and editors, n. (ed.), (2001), *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Allyn & Bacon. Boston, MA (Pearson Education Group).
- Andrade H. and Du Y., (2005), Student perspectives on rubric-referenced assessment, *Practical Assessment, Research & Evaluation*, **10**, (note—this is an electronic journal and there are no page numbers).
- Andrade H. G., (2000), Using rubrics to promote thinking and learning, *Educational Leadership*, **57**, 13–18.
- Andrade H. G., (2005), Teaching with rubrics: The good, the bad, and the ugly, *College Teaching*, **53**, 27–30.
- Bissell A. N. and Lemons P. P., (2006), A new method for assessing critical thinking in the classroom, *BioScience*, **56**, 66–72.
- Black P. and Wiliam D., (1998), Assessment and classroom learning, *Assessment in Education: Principles, Policy & Practice*, **5**, 7.
- Bloom B. S., (1956), *Taxonomy of educational objectives: The classification of educational goals*. New York: McKay.
- Bullough R. V. and Pinnegar S. E., (2007), Thinking about the thinking about self-study. In J. Loughran, M. H. Hamilton, V. K. LaBoskey, and T. Russell (ed.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Dordrecht, The Netherlands: Springer.
- Brown J. S., Collins A. and Duguid P., (1989), Situated cognition and the culture of learning, *Educational Researcher*, **18**, 32–42.
- Burke K. A., Greenbowe T. J. and Hand B. M., (2006), Implementing the science writing heuristic in the chemistry laboratory, *Journal of Chemical Education*, **83**, 1032.
- Carless D., (2007), Conceptualizing pre-emptive formative assessment, *Assessment in Education*, **14**, 171–184.
- Earl L. M., (2003), Assessment as learning: Using classroom assessment to maximize student learning. Thousand Oaks, CA: Corwin Press.
- Ennis R. H., (1985), A logical basis for measuring critical thinking skills, *Educational Leadership*, **43**, 44–48.
- Fink L. D., (2003), *Creating Significant Learning Experiences: An integrated approach to designing college courses*. San Francisco: Jossey-Bass.
- Fitch G. K., (2007), A rubric for assessing a student's ability to use the light microscope, *The American Biology Teacher*, **69**, 211–214.
- Good R. and Shymansky J., (2001), Nature-of-science literacy in “benchmarks” and “standards”: Post-modern/relativist or modern/realist?, *Science and Education*, **10**, 173–185.
- Hafner J. and Hafner P., (2003), Quantitative analysis of the rubric as an assessment tool: An empirical study of student peer-group rating, *International Journal of Science Education*, **25**, 1509–1528.
- Halonen J. S., Bosack T., Clay S., McCarthy M., Dunn D. S., Hill Iv G. W., et al., (2003), A rubric for learning, teaching, and assessing scientific inquiry in psychology, *Teaching of Psychology*, **30**, 196–208.

- Halpern D., (1996), *Thought & knowledge: An introduction to critical thinking*. Mahwah, NJ: Lawrence Erlbaum.
- Halpern D., (2001), Assessing the effectiveness of critical thinking instruction, *Journal of General Education*, **54**, 238–254.
- Hand B., Prain V., Lawrence C. and Yore L. D., (1999), A writing in science framework designed to enhance science literacy, *International Journal of Science Education*, **21**, 1021–1035.
- Hunter A.-B., Laursen S. L. and Seymour E., (2007), Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development, *Science Education*, **91**, 36–74.
- Jacobs C., (2004), Critical thinking in the chemistry classroom and beyond, *Journal of Chemical Education*, **81**, 1216–1223.
- Jiménez-Alexandre M. P., Bugallo Rodríguez A. and Duschl R. A., (2000), “Doing the lesson” or “doing science”: Argument in high school genetics, *Science Education*, **84**, 757–792.
- Jonsson A. and Svingby G., (2007), The use of scoring rubrics: Reliability, validity and educational consequences, *Educational Research Review*, **2**, 130–144.
- Kardash C. M., (2000), Evaluation of an undergraduate research experience: Perceptions of undergraduate interns and their faculty mentors, *Journal of Educational Psychology*, **92**, 191–201.
- Lave J. and Wenger E., (1991), *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.
- Lederman N. G., (1992), Students' and teachers' conceptions of the nature of science: A review of the research, *Journal of Research in Science Teaching*, **29**, 331–359.
- Lederman N. G., (2007), Nature of science: Past, present and future. In S. K. Abell & N. G. Lederman (ed.), *Handbook of research on science education*. New York: Routledge.
- Liang L. L., Chen S., Chen X., Kaya O. N., Adams A. D., Macklin M., et al., (2006), *Student understanding of science and scientific inquiry: revision and further validation of an assessment instrument*. Paper presented at the Annual Conference of the National Association for Research in Science Teaching (NARST).
- Loughran J. J., (2007), A history and context of self-study of teaching and teacher education practices. In J. Loughran, M. H. Hamilton, V. K. LaBoskey, and T. Russell (ed.), *International Handbook of Self-Study of Teaching and Teacher Education Practices*. Dordrecht, The Netherlands: Springer.
- Maki P., (2004a), Assessing for learning: Building a sustainable commitment across the institution. Sterling, VA: American Association for Higher Education.
- Maki P., (2004b), Reaching consensus about criteria and standards of judgment, Chapter 5. Retrieved from http://www.uri.edu/assessment/media/public/page_files/resources/maki/Maki_Consensus_about_Criteria_and_Standards.pdf.
- Marchlewicz S. C. and Wink D. J., (2011), Using the activity model of inquiry to enhance general chemistry students' understanding of nature of science, *Journal of Chemical Education*, **88**, 1041–1047.
- McComas W. F., (1998), Principle elements of the nature of science: Dispelling the myths. In W. F. McComas (Ed.), *The nature of science in science education*. Kulwer: Springer.
- McCray R. A., DeHaan R. L. and Schuck J. A., (ed.), (2003), *Improving Undergraduate Instruction in Science, Technology, Engineering, and Mathematics. Report of a Workshop*. Washington D.C.: National Academies Press.
- Mertler C., (2001), Designing scoring rubrics for your classroom, *Practical Assessment, Research & Evaluation*, **7**.
- Moni R. W., Beswick E. and Moni K. B., (2005), Using student feedback to construct an assessment rubric for a concept map in physiology, *Advances in Physiology Education*, **29**, 197–203.
- NRC., (1996), *National Science Education Standards*. Washington, D.C.: National Academy Press.
- NRC., (2011), *A Framework for K-12 Science Education: Practices, Cross-Cutting Concepts, and Core Ideas*. Washington, D.C.: National Academy Press.
- Pascarella E. T. and Terenzini P. T., (2005), *How college affects students: A third decade of research*. San Francisco, CA: Jossey-Bass.
- Popham W. J., (1997), What's Wrong—and What's Right—with rubrics, *Educational Leadership*, **55**, 72–75.
- Quitadamo I. J., Faiola C. L., Johnson J. E. and Kurtz M. J., (2008), Community-based inquiry improves critical thinking in general education biology, *CBE-Life Sciences Education*, **7**, 327–337.
- Reddy Y. M. and Andrade H., (2009), A review of rubric use in higher education, *Assessment & Evaluation in Higher Education*, **35**, 435–448.
- Ruiz Primo M. A., Li M., Ayala C. and Shavelson R. J., (2004), Evaluating students' science notebooks as an assessment tool, *International Journal of Science Education*, **26**, 1477–1506.
- Russell C. B. and Weaver G. C., (2011), A comparative study of traditional, inquiry-based, and research-based laboratory curricula: Impacts on understanding of the nature of science, *Chemistry Education Research and Practice*, **12**, 57–67.
- Schwartz R. S., Lederman N. G. and Crawford B. A., (2004), Developing views of nature of science in an authentic context: An explicit approach to bridging the gap between nature of science and scientific inquiry, *Science Education*, **88**, 610–645.
- Singer J., Marx R. W., Krajcik, J. and Clay Chambers J., (2000), Constructing extended inquiry projects: Curriculum materials for science education reform, *Educational Psychologist*, **35**, 165–178.
- Tariq V. N., Stefani L. A. J., Butcher A. C. and Heylings D. J. A., (1998), Developing a new approach to the assessment of project work, *Assessment & Evaluation in Higher Education*, **23**, 221–240.
- Timmerman B. E. C., Strickland D. C., Johnson R. L. and Payne J. R., (2010), Development of a ‘universal’ rubric for assessing undergraduates' scientific reasoning skills using scientific writing, *Assessment & Evaluation in Higher Education*, 1–39.
- Wong S. L. and Hodson D., (2009), From the horse's mouth: What scientists say about scientific investigation and scientific knowledge, *Science Education*, **93**, 109–130.
- Zoller U., (1993), Are lecture and learning compatible? Maybe for LOCS: Unlikely for HOCS, *Journal of Chemical Education*, **70**, 195–197.
- Zoller U., (1999), Scaling-up of higher-order cognitive skills-oriented college chemistry teaching: An action-oriented research, *Journal of Research in Science Teaching*, **36**, 583–596.
- Zoller U., (2001), The challenge for environmental chemistry educators, *Environmental Science and Pollution Research*, **8**, 1–4.