The Geological Society of America
Special Paper 461
2009

The evaluation of field course experiences: A framework for development, improvement, and reporting

Eric J. Pyle
Department of Geology and Environmental Science, James Madison University, MSC 6903, Harrisonburg, Virginia 22807, USA

ABSTRACT

There is little argument that field course experiences are both complex and unique in the range of learning experiences provided to students. Conversely, they offer logistical and cost challenges that might cause one to question whether they provide a sufficient cost-benefit ratio to warrant continuation, particularly in a climate where resources have become scarce. In such a climate, it is important to have on hand rigorous data that support assertions of learning effectiveness. Many of the data supporting the evaluation of field course experiences can come from an analysis of assessments of student performance relative to course goals, but these data alone may not provide sufficient support. A close examination of faculty actions relative to student learning outcomes, as well as a research-based analysis of course curricula designed to best support student learning, can provide two additional sources of data. When used in concert with student assessment data, evaluative success can be triangulated. A consistent set of tools in this evaluative framework also provides information on specific areas for maximizing student learning. This chapter outlines such a set of tools, using a specific field course experience that is in transition as a model. Pilot data collected within the existing field course experience structure are discussed in a manner that informs the development of performance assessments, instructional actions, and curricular organization. Using data derived from these sources, evaluations of field course experiences can be used to better articulate the cost-benefit ratio in terms of student learning in the cognitive, affective, and psychomotor domains.

INTRODUCTION

There is little doubt that field camp experiences, or field course experiences, are intensive of financial, faculty, and material resources. As costs have risen, it is not an unreasonable question to ask if the investment is worth the outcome. A cursory review of the intended outcomes of field course experiences, as posted online (Baker, 2006; King, 2009) provides a generally consistent view that field course experiences serve to hone students' skills, prepare them for the workplace, allow them to apply classroom-based learning to real situations, serve as a capstone learning experience, or immerse them in the conventions and expectations of professional geoscientists. These outcomes and values are universally valued within geoscience departments (Baker, 2006). However, outside of geoscience departments, the challenge is to provide administrators with a justification for the resource-intensive nature of field course experiences, especially in a climate of budget shortfalls and (relative to other departments) lower enrollments in geoscience programs. Academic freedom lasts right up to financial exigency, and then the need for clear justification becomes paramount.

Field Course Experience

There is a considerable body of research literature focusing on the nature of effective science learning experiences that

Pyle, E.J., 2009, The evaluation of field course experiences: A framework for development, improvement, and reporting, *in* Whitmeyer, S.J., Mogk, D.W., and Pyle, E.J., eds., Field Geology Education: Historical Perspectives and Modern Approaches: Geological Society of America Special Paper 461, p. 341–356, doi: 10.1130/2009.2461(26). For permission to copy, contact editing@geosociety.org. ©2009 The Geological Society of America. All rights reserved.

indicates learning is constructed by students as facilitated by their instructors and instructional environments (Resnick, 1983; Anderson, 1987; Mestre and Cocking, 2000; Bybee, 2002). This concept is not alien to the geosciences, as was suggested by T.C. Chamberlin. In his mind, an important consideration in Earth inquiries is that students should create "by [their] own effort an independent assemblage of truth" (Chamberlin, 1896, p. 848). What becomes apparent early in any inquiry in the earth sciences is that the questions are often based on incomplete information about complex, interactive, and (ultimately) uncontrollable events, and thus, these questions defy simple or discrete explanation through any single pathway of inquiry (Ault, 1998; Frodeman, 1995). Getting lost in the complexity is easy, so when instructors fall back on questions that are trivial or limited to confirmation of previous results, it is perhaps merely defensive and "safe" in instructional situations. Given the ambiguity and uncontrollability of geoscience phenomena, the conservative approach would favor instruction that demonstrates effectiveness in situations unsuited and not supportive of field course experiences, and yet students are placed squarely in these (at least to their perspective) complex and ambiguous situations. The complexity that is inherent in a field course experience is a unique learning experience that solidifies the knowledge, skills, and dispositions for professional growth (Stokes and Boyle, this volume).

Keeping the complexity of the field course learning experience in mind, an evaluation framework that seeks to document the value-added nature of field course experiences, as well as a favorable cost-benefit ratio, should provide more information than student performance alone. Furthermore, evaluation should work complementarily with development, such that one informs the other. This manuscript examines the various aspects of student learning that could and should be examined in the context of a field course experience, the ways faculty interact with students to promote this learning, and the elements the curriculum should include to support the desired learning. Using the case of a field course experience in a developmental transition, the relationship of students, faculty, and curriculum to the field-based knowledge, skills, and dispositions that are developed in a field course experience are considered. Specific questions to be addressed by this paper are: (1) What should student performance assessment include to meet learning outcomes in the field course experience? (2) How can faculty involvement be documented that supports these learning outcomes? (3) What elements should be considered when designing instruction that can be employed by faculty to best promote student learning? Based on data collected during a recent field course experience, these data will be used to inform a set of tools that can be readily used by other field course experience providers to evaluate their own offerings for internal and external audiences. Furthermore, the definition of this evaluation framework sets the stage for implementation in the first, post-transitional, offering of the course. Through such a comprehensive approach to evaluation, the justification for field course experiences should be evident, not just to geoscientists, but also to academic administrators.

Assessment versus Evaluation

It has been said that if one does not like evaluation, then education is the wrong business for them to be in. The terms "evaluation" and "assessment" tend to be used interchangeably in common practice, but for the purposes of this chapter, each will have a specific definition. Ebert-May (1998) defines assessment as "data collection with a purpose," while Frechtling (2002, p. 3) defines evaluation as "the systematic investigation of the merit or worth of an object."

Assessment involves comparing information gathered from subjects relative to some established goal or objective (Kizlik, 2009). These goals, objectives, or outcomes are set in advance, and should be clear to both instructors and students. Through the use of a valid assessment that yields consistent results, the impact of instruction on students can be determined by the extent to which they have met or demonstrated these established goals. Thus, there is no "good" or "poor" as a part of assessment, only the difference between student performance on the task and the expectations established by the goal. Arguably, there is more familiarity with tasks tied to either cognitive (knowledge)-based objectives or, to a lesser extent, those tied to psychomotor (skill)-based objectives. Affective outcomes that define dispositions or habits of mind are often overlooked because these outcomes are often more implicit and more difficult to measure.

Evaluation allows us to establish and communicate the worth of an activity to internal and external audiences (Kizlik, 2009). To internal audiences, this worth can be determined by the extent to which decisions of instructional approaches, arrangements, organization, etc., are effective in aiding students to reach the desired outcomes. Such worth is determined by, but not limited to, the assessment data that are routinely collected. This, in essence establishes (or not) the validity of such choices. With respect to external audiences, "worth" can be determined by cost-effectiveness of effort relative to students meeting expectations, or through the establishment of the appropriateness of experiences to an overall curriculum model or larger set of expectations. These determinations become statements of "value-addedness" to student preparedness for future professional roles.

Field Course Experience Outcomes

As is implied in describing the general importance of field course experiences, the geosciences have a unique set of conventions and methodologies, supported by both general as well as specialized philosophies of science (Kitts, 1977; Frodeman, 2003). To experts in the field, these conventions and methodologies are largely transparent; they are just how things are done. However, as Gardner (1993) pointed out, once one becomes an expert, it is difficult to remember how it is to *not* be an expert and not know. Therefore, in considering the preparation and professional development of future geology professionals, it is useful to have a framework to "remember" how geoscientists come to know, act, and feel within their practice (see also Chi et al., 1981; Bransford et al., 2000).

Explicitly, then, field course experiences are intended to reinforce the skills of a geologist at an early precareer stage:

> Field camp is a tradition in the education of a geologist. It is an intensive course that applies classroom and laboratory training to solving geological problems in the field. Skills developed during field camp typically include: collection of geologic data, constructing a measured section, interpreting geologic structures, and geologic mapping. (King, 2009)

To view contemporary field course experiences relative to one another, Geology.com maintains a list of currently available field course experiences (King, 2009), as does American Geological Institute (AGI) (Baker, 2006). Sadly, relatively few field course experiences provide explicit goals and objectives as a part of the general description, nor do they often provide syllabi from which information may be drawn. From the available, explicit information, the following points of commonality are seen:

1. Recall or comprehension of facts is secondary to actually utilizing and applying facts, in that the "facts" are assumed to have been mastered by (or are at least familiar to) students, whereas the use of this knowledge through data analysis and synthesis of solutions is much more prominent.

2. Participants learn the use and application of equipment, tools, and techniques in field geology, focusing on the skill set necessary to function as an entry-level professional geologist.

3. Participants develop the habits of mind that govern the application of those knowledge and skills with integrity and attention to detail, valuing the conventions, techniques, and communication skills that make geology a rigorous science.

4. It is important to see each of these goals expressed in a variety of contexts, such that students' development as geologists is enriched by their exposure to a variety of geologically interesting contexts.

Many of these aspects of field course experiences are expressed as general goals rather than as specific objectives. As a result, they form the core statements that can be used to formulate not only specific objectives used in assessment, but also general questions for the evaluation of field course experiences. However, to do so, the knowledge, skills, and dispositions to be learned in field course experiences must be made explicit by instructors to students and external audiences. (Please see Appendix 1 for a sample of field course experience outcomes.)

A CASE STUDY: JAMES MADISON UNIVERSITY'S FIELD COURSE EXPERIENCE

The Department of Geology and Environmental Science at James Madison University (JMU) has operated a 6 wk geology field camp in the Connemara Peninsula of Western Ireland since 2005. This field course is conducted in cooperation with the National University of Ireland–Galway, and was originally developed by Boston University. The explicit description of the field course experience is described in the syllabus as:

> After completing the field course, you will be qualified to work for an industrial, governmental, or academic employer who needs you to make your own way to an isolated village in a foreign country, assess the local geology, natural resources, natural hazards, environmental conditions, etc., write a project report, draft a publishable map, generate a data base, and return home safely. The main objective is for you to become confident at scientific observation, interpretation, and solution of geological problems in the field. You will learn to recognize and interpret a wide variety of rock types, structures, and geomorphic features. We will place emphasis on methods of map-making, data recording, and report preparation. Projects from one to five days duration will be conducted in well-exposed igneous, metamorphic, and sedimentary rocks, ranging in age from Precambrian through Quaternary and correlative to rocks and sediments of the northern Appalachians.

The 2008 offering of the course was a transitional year because the administration passed fully over to JMU, while several new faculty members were added to the course. Much of the course structure and many of the exercises remained unchanged, although they were sequenced in a manner reflective of available faculty expertise. This created an opportunity to explore the development of an evaluation framework for the field course, such that the learning value and adherence to goals could be documented in a comprehensive fashion that would eventually not only justify the expense of the course, but also provide information on the efficacy of the particular scope and sequence of learning activities that make up the field course experience. The 2008 data collection, described herein, was not intended to provide these specific answers, but to generate ideas for a framework to be employed in future offerings for evaluation and continued development.

Several primary sources of data were used during the 2008 course offering. First, each of the 29 students were asked to complete a brief questionnaire, outlining not only their prior course experience, but also their personal level of confidence with respect to that course, scored on a 0–5 scale, 5 being "very confident." These two pieces of data were designed to capture crude information that could inform the development of evaluation questions on student preconceptions and metacognition. Fifteen students came from James Madison University, eight came from Boston University, and the remainder came from other institutions. Students were also asked about their prior use of geologic tools, such as compasses and global positioning system (GPS) units. The results of this questionnaire are found in Figures 1 and 2 below.

Students indicated prior experience with traditional coursework in geology, including physical, historical, and structural geology, as well as mineralogy and petrology. Fewer students had taken stratigraphy and geomorphology, and fewer still had previously taken specialized courses such as tectonics, paleontology, and sedimentology. Only a few students had taken environmentally oriented courses. Interestingly, students expressed a

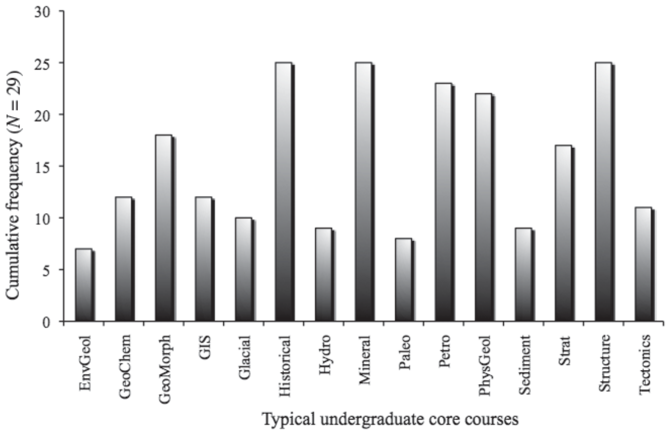


Figure 1. Frequency of student course experience in prior geology coursework common to undergraduate geology programs; *N* = 29. GIS—geographic information systems.

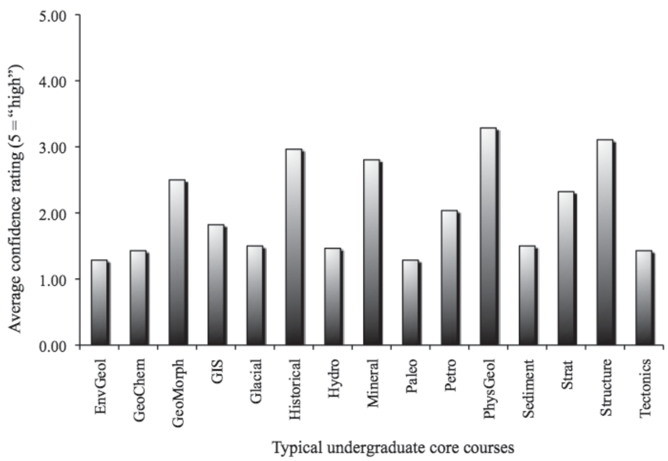


Figure 2. Mean student confidence level with mastery in prior geology coursework common to undergraduate geology programs: 1—low confidence, 5—high confidence. GIS—geographic information systems.

confidence range that largely paralleled their prior experience, although not at a level that would reflect a belief in personal mastery of the material, as confidence never exceeded 3.5.

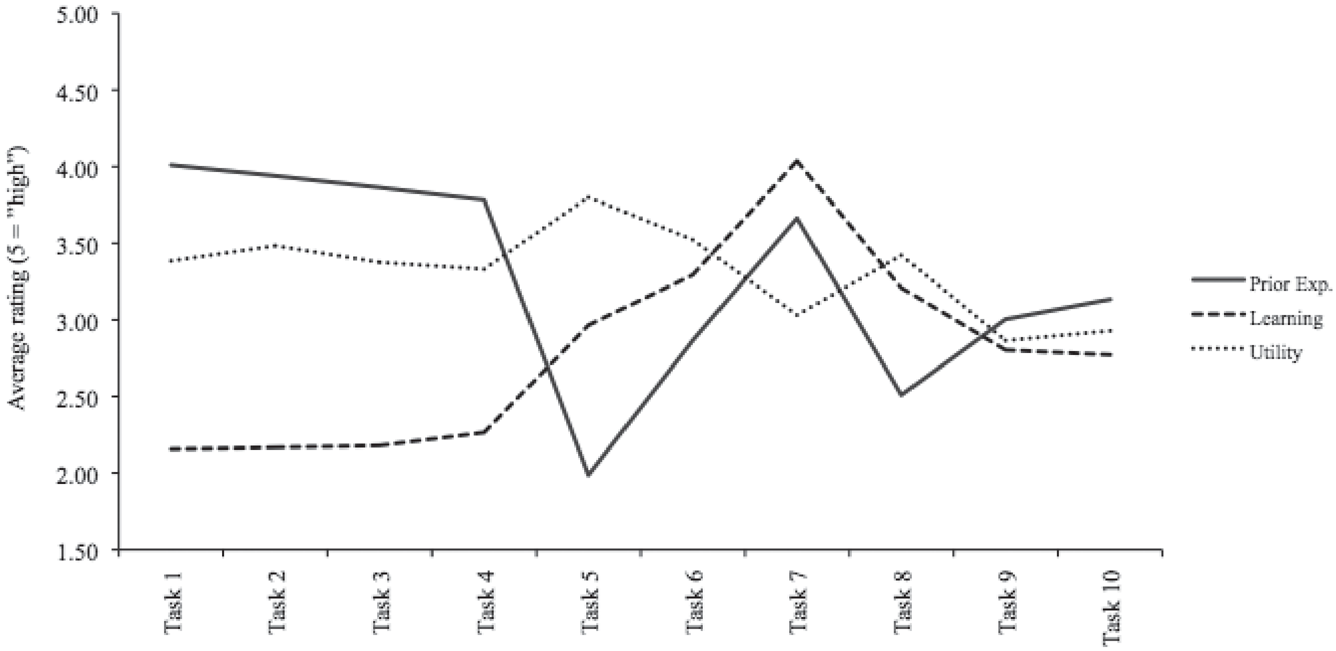
During the progress of the course, students were also asked for responses on the specific exercises, reflecting on their experience with course exercises, on a 1–5 Likert scale (5 being "high," "great," or "very useful"). These were administered at the end of week 3 and again at the end of the field course experience (week 6). Students were asked about their prior experience with the material that made up the exercise, their perceived level of learning from the experience, and their perceptions of the utility of that learning. These data were plotted across the course sequence and are summarized in Figure 3. Additional narrative data were also collected for each exercise, drawing from open (anonymous) comments as well as observational notes, personal reflections, and brief post–field course experience interviews.

It was expected that the level of prior experience with the material at each site would start relatively low and then increase. Instead, it started relatively high, showed variation in the middle of the camp, and then returned to a lower level than the start. It was also expected that the students' perception of learning after each exercise might start high and would show an increase over time, as the range of experiences increased. Overall, the level of learning did increase, but in a nonuniform manner, starting at a low level, peaking near week 4, and then decreasing. Finally, students perceived utility of the exercises were expected to start low and then increase. Instead, student perceptions of the utility of exercise started relatively high and decreased slightly as the course progress.

These student reports are quantitative, but because they are self-reports and largely categorical data, they are of limited value in an evaluative sense. Furthermore, the written comments are anecdotal, reflecting specific episodes or narrow perspectives on interactions among faculty, students, and the curriculum. Thus, the questions that students were asked provide a limited basis for assessing skills and dispositions, but they do not comprise a true rubric for determining skills and dispositions changes. As a result, it was agreed that the data collected during the 2008 field course offering provided an appropriate basis for student assessment, but it was an incomplete data set for general evaluative purposes. The instruments were not constructed with broad generalizability in mind, nor were they necessarily meant to demonstrate reliability across course offerings. Rather, they were intended to provide a general student evaluation of instruction, with at least face validity and limited content validity. Taken as generative data (Goetz and LeCompte, 1984), however, they suggested strands that form the basis for the evaluation questions stated in the introduction.

Solid inferences based on these results are difficult to make, but given the exploratory nature of this investigation, the results are suggestive of a number of commonalities that invite more detailed study. For example, it would appear from the quantitative data that the sequence of exercises could perhaps have been better matched to the particular set of students. There were little data to support the representativeness of this particular population of students, either in their prior knowledge, skills, or their capacity for professional self-awareness. The sensitivity of the instrumentation is insufficient at this time, but it has been adjusted for the next offering of the course. Already, the nature of the course has been restructured, such that student preconceptions and mastery of field-based inquiry are directed toward their interest in either general geologic problems or environmental techniques, with an aim to promoting a professional self-identity.

The results underscore the future utility of the data in an overall evaluation framework, one that is demonstrably linked to goals. The documentation of these student data tied to their performance is a necessary component of additional data to sup-



Task	Week	Nature
1	1	Individual geologic map
2	2	Group glacial features mapping
3	2	Group metamorphic bedrock map
4	3	Group stream environmental analysis
5	3	Group digital mapping
6	4	Large-scale geologic structures memoir
7	4	Independent mapping
8	5	Regional geology memoir
9	5	Tectonic environments memoir
10	6	Individual karst terrain mapping

Figure 3. Changes in student reports of prior experience, perceived learning, and perceived utility of exercises across the span of the 2008 James Madison University field course.

port an informed evaluation. A more sensitive means is needed to determine the ways in which students grow toward meeting the outcomes of the field course experience. The manner in which faculty in general promoted this growth through their interactions with students or instructional decisions is not well documented in the current framework. Another aspect that is not well documented is the way in which the curriculum was designed to have students meet explicit and implicit course outcomes. The remainder of this manuscript thus defines not only a way that sensitivity of student assessments can be enhanced, but also ways in which faculty engagement can be documented within a curricular framework that research on science learning has demonstrated to be effective in promoting deep student learning. Plans for future offerings of the JMU field course experience are used as examples in each of these contexts.

STUDENT ASSESSMENT

The available literature on student assessment in field course experiences is focused to a large extent on the cognitive outcomes, identifying the content of what should be learned in field course experiences by different audiences (Anderson and Miskimins, 2006) or comparing field and laboratory components of a student's program experience (Noll, 2003). Measures of student learning are largely quantitative but limited to objective test or pre- to postexperience comparisons. There is an implicit attention to issues of skill and professional mind-sets, but these are not measured in detail in these studies. Hughes and Boyle (2005) argued for forms of assessment unique to the earth sciences and made a clear distinction between class work, laboratory work, and fieldwork, as each requires distinct approaches to assessment. Furthermore, the arguments for establishing the validity and reliability of assessments are strong (Butler, 2008), whether considering fieldwork in both class and residential program contests. While not specifically stipulated by Butler (2008), these assessments can provide useful program evaluation data.

While field course experience learning in the cognitive domain is well represented, there is less representation of student growth in the affective or psychomotor domains, making these forms of data normally unavailable for program evaluation. Boyle and his colleagues (2007), however, provided comprehensive measures of student affect as a part of fieldwork, as do Stokes and her colleagues (this volume), concluding that while there are increases in positive student feeling toward fieldwork after the experience, there are also suggestions that affect plays a greater role in professional dispositions than had previously been documented in the geoscience education literature. Interestingly, most of the information on student learning of skills and dispositions comes from the geoscience education literature that focuses on earth science teachers. Since professional development programs for earth science teachers are often externally funded through grants, there is a need for comprehensive evaluation in order to ensure that the projects have a positive impact on teachers, and not just the teacher participating, but also on their students. In order to enhance the experience of the teachers, they are often engaged in authentic research experiences involving considerable amounts of fieldwork. Measures of teacher skills and dispositions related to the practice of geology are well documented by Huntoon and her colleagues (2001), O'Neal (2003), and Hemler and Repine (2006). In each of these projects, multiple and varied forms of assessment data were used, including recognized forms in geology such as maps, field notes, and cross sections. They also expanded the assessment repertoire to include teacher artifacts such as concept maps, lesson plans, journals, and constructed responses. These additional forms of data were used to triangulate gains in knowledge, skills, and dispositions in these studies.

Techniques of Assessment That Reflect the Structure of the Geosciences

Every assessment, regardless of its purpose, rests on three pillars: (1) a model of the way students represent knowledge and develop competence in the subject domain, (2) tasks or situations that allow one to observe students' performance, and (3) an interpretation method for drawing inferences from the performance evidence thus obtained. In the context of large-scale assessment, the interpretation method is usually a statistical model that characterizes expected data patterns given varying levels of student competence. In less formal assessment, the interpretation is often made by an instructor using an intuitive or qualitative insight, rather than statistics, focused less on a determinative and more on a developmental purpose (Atkin and Coffey, 2003). If then, assessment is to be effective, it needs to be demonstrably tied to learning goals, whether they are reflective of knowledge, skills, or dispositions (Fox and Hackerman, 2003). The difficulty for earth science instruction lies in the intrinsically interdisciplinary nature the geosciences (Hughes and Boyle, 2005), and crafting not only instruction but also assessment to represent this format and, thus, attain validity of the assessment.

An understanding of the purpose and format of an assessment is a prerequisite to ensuring the consistency and reliability of both administration and interpretation of assessment data. Furthermore, the complexities of the contexts of earth science instruction, whether in class, the laboratory, or in the field, demand that assessment be explicit in reflecting these different settings and intended uses. Assessments can be seen as *formative*, in which the level of student achievement in particular objectives is communicated back to students in order to promote continued growth toward mastery, but also to faculty in order to indicate course corrections. Assessments can also be seen as s*ummative*, in that they are used to provide a final determination of student achievement relative to the goals of instruction. These data are also used for comparison, group analysis, and external reporting. Given these formats for assessment, it is necessary to parse the task into elements reflective of knowledge, skills, and dispositions. The following is a brief summary of the ways in which assessment elements in each of these areas can be designed, based first on the literature and then defined with field course experience–specific task suggestions.

Knowledge

Decades of research on student learning and instructional design have produced a variety of taxonomies that are useful for a systematic means of parsing knowledge for both instruction and assessment. Perhaps the best known is Bloom's cognitive taxonomy, which is discussed in a variety of sources (Bloom, 1956; Trowbridge, Bybee, and Powell, 2004). In developing objectives in the cognitive, or for that matter each, domain, the challenge is to frame it around an active, measureable verb, stating both the task that is expected of students as well as the criteria that indicate student mastery of that objective (Chiappetta and Koballa, 2006). Using this taxonomy, many familiar field course experience tasks are provided with clear, measureable definitions that communicate internally as well as externally. Application of this taxonomy to field course experiences is suggested in Table 1.

These elements have become increasingly important in assessment of students, but one should view this use with some caution. It is relatively easy to devise assessment items of high validity and reliability at the first two levels, the lower-order thinking skills, than it is for the latter four, or higher-order thinking skills. Nevertheless, this taxonomy is best used in the creation of instructional objectives that many can agree upon as important

TABLE 1. COGNITIVE TAXONOMIC ELEMENTS REFERENCED WITH RESPECT TO POTENTIAL FIELD COURSE EXPERIENCE TASKS OR EXPECTATIONS OF STUDENTS, WITH ACTIVE VERBS TO FRAME THE OBJECTIVE

Cognitive level	Sample verbs	Earth science concept
Knowledge	Define, describe, identify	Rock texture, RFM identification
Comprehension	Interpolate, estimate, predict	Draw contour lines from elevation data
Application	Compute, modify, relate, use	Graph a topographic cross section
Analysis	Diagram, divide, infer	Plot fold axis on a map
Synthesis	Arrange, generate, design	Construct a geologic map from field data
Evaluation	Contrast, interpret, appraise	Assess landslide hazards from map data

RFM—rock-forming mineral.

for students to have mastered in order to be successful in their employment or in graduate school.

Psychomotor

The sciences, when practiced for the generation and verification of new knowledge, rely on not just the application of discrete knowledge, but also on the application of set of specialized skills. These skills are typically referred to as *psychomotor*, indicating that there is a brain-body connection of some definable nature. There are several models of psychomotor taxonomies (e.g., Simpson, 1972), but the work of Dave (1975) matches well to field course experience tasks and supports the development of measureable objectives. Like the cognitive taxonomy described previously, they can be ordered in increasing level of difficulty, as in Table 2.

One aspect that should be evident from this limited introduction to the psychomotor domain is that the geosciences are of special concern. For example, the observational skills required in the geosciences necessitate attention to the details of a phenomenon as well as the larger context. To understand a flood in a cognitive manner requires observing with precision the details of a stretch of streambed (shape, sediment load, etc.) as well as the larger context (recognizing and measuring the floodplain from contour maps, measuring changes in flow rate, etc.). In addition, many of these observations rely heavily on the visual domain, both in pattern recognition as well as communication of ideas, such that written descriptions and verbal presentations become an adjunct to diagrams, charts, and illustrations, rather than the text as the leader. This is a complex skill that must be cultivated among students if they are to function with a high level of content-related skill.

Dispositions

The third domain to consider in the preparation of geoscientists deals with the starting point in thinking and acting, namely, one's dispositions and habits of mind. Arguably, these starting points are first governed by the affective domain, which is concerned largely with feelings and emotions, but they are not limited this area. Instead, they drive the basic template of a student's approach to a problem or unique situation, and they strongly influence attitudes and potential actions (Azjen and Fishbein, 1980). Like knowledge and skills, affective dimensions can be taxonomically arranged (Krathwohl et al., 1973), as in Table 3.

Among the three domains discussed here, dispositions and affect are perhaps the most difficult to measure or assess. More importantly, they are likely the objectives most difficult to explain to those outside of the geosciences, or for that matter, any science. However, they are also clearly a part of the covert curriculum, and few instructors would not attach some value or professional satisfaction to students clearly attaining these objectives.

The knowledge, skills, and dispositions outlined here are a first step in representing the structure of the discipline in instruction and assessment. Returning to the structure of the discipline, assessment items or tasks can be built around: (1) knowledge-based representations, as distinct from "knowledge" as beliefs described previously; (2) lexical representations of terminology specific to context; and (3) prototypes or exemplars, which are in part model or graphical representations of phenomena (Smith, 1995; Lawrence and Margolis, 1999; Murphy 2002; Sibley, 2005). More specific task/item examples are provided in Table 4.

TABLE 2. PSYCHOMOTOR TAXONOMIC ELEMENTS USEFUL TO FIELD COURSE INSTRUCTION AND ASSESSMENT

Psychomotor level	Sample indicators	Earth science action
Imitation	Crude reproduction of action based on observation and minimal practice	Determination of mineral sample physical properties, such as hardness, streak, or observing cleavage
Manipulation	Performance from instruction with attention to form	Measurement and data encoding using a Brunton compass or Jacob's staff
Precision	Accuracy, proportion, and exactness in performance, with minimal error	Collection of physical and chemical data at several points along a stream
Articulation	Coordinating a series of acts with harmony and consistency	Map generation from a series of station measurements, plotted on a base map
Naturalization	Smooth, natural performance with minimum of psychic energy	Generation of finished maps that reflect multiple layers of data collection and procedures and coordinate well with field notes and diagrams

TABLE 3. AFFECTIVE TAXONOMIC ELEMENTS THAT SHOULD INFLUENCE FIELD COURSE INSTRUCTIONAL DESIGN AND STUDENT ASSESSMENT

Dispositional level	Sample indicators	Earth science action
Receiving	Follow directions, locate, identify	Following along with a field trip guidebook as a part of a field trip
Responding	Complete assigned tasks at or above level required, or for self-satisfaction	Once a local geologic map has been studied, seeking out a regional geologic map to see larger context
Valuing	Accept, prefer, and commit to scientific values	More than one measurement of a particular parameter is sought in each location
Organizing	Personal values are brought into line with scientific values	Each field investigation is approached with a set of questions framed on methodologies and possible outcomes
Characterizing	Lifestyle adoption indicative of a preference for scientific values	Active seeking of communications with other students and faculty on geological issues

TABLE 4. FIELD COURSE EXPERIENCE ASSESSMENT ELEMENTS, PROVIDING A BASIS FOR RUBRIC DEVELOPMENT AND ASSESSMENT PLANNING

Structure	Sample task within field course	Rationale from knowledge, skills, and dispositions
Knowledge-based	Based on observations of current stream conditions and local erosion and sedimentation patterns, making a prediction of how the stream changes when flow reaches flood conditions.	Knowledge—distinguishing beliefs from prior knowledge, applied to novel situation Skills—measurements and observational descriptions of setting Dispositions—use of more than one parameter in making the prediction
Lexical	Correctly applying terminology in a lithologic description using texture, and mineralogy, and internal features or structures.	Knowledge—recall and appropriate application of terminology Skills—effectively communicating descriptions in written or oral form Dispositions—using a variety of descriptive terms in a manner that reflects possible contexts
Prototype	Constructing an accurate cross section from a map, or distinguishing the correct cross section from distracters, stating reasons for rejection.	Knowledge—synthesizing an analogy representing the distribution and orientation of materials Skills—drafting a cross section with consistency of measurement, to scale, from the map Dispositions—cross section contains all necessary detail, drafted in a manner that communicates clearly the interpretations drawn

In application, these elements provide not just summative assessment data, but they can also serve to generate formative assessment data, teasing out student preconceptions when designing or modifying instruction, selecting particular prior learning that can be built upon or that needs particular attention in subsequent instruction. When used as a form of embedded assessment, they can provide direct support to student-led inquiry, such that their application of professional skill sets is evident. Finally, they serve as a jumping-off point for deeper self-reflection and professional self-awareness, providing currency and a real-world focus that can be directly applied to the world outside of class. If these tasks are to support student learning, they should be constructed in such a manner, so that students feel they have the latitude to pursue novel solutions that may deviate from conventional solutions (Hughes and Boyle, 2005).

Based on the data from the 2008 offering of the JMU field course experience, considered in light of the assessment elements discussed here, a new set of rubric elements has been developed for field course experience tasks. It is intended to be drawn on as a bank of statements, to the extent that a given task may be knowledge, lexical, or prototype in nature and thus require a specialized framework for determining student mastery of learning goals. These statements and mastery descriptors are offered in Appendix 1, but sample elements to be employed in the 2009 offering of the JMU field course experience are offered in Table 5.

Astin and his colleagues (1996) argued that student assessment needs to adhere to several characteristics in order to contribute to meaningful evaluation. In the context of field course experience evaluation, assessments should have the following characteristics:

1. Assessments should embody a vision for the most valuable kinds of learning—Knowledge, skills, and dispositions that are important for an entry-level professional geologist should not only be part of the assessment techniques, but these assessments should be evident to students, faculty, and external audiences.

2. Assessments should be multidimensional, integrated with instruction, and reflect performance over time—Assessments should be as much of the overall developmental sequence as instruction, beginning with more general ideas and moving toward specific performances.

3. Assessments are best when tied to clear expectations and purpose—To the extent that students know clearly what they are

TABLE 5. APPLICATION OF ASSESSMENT STRUCTURE ELEMENTS IN THE DEVELOPMENT OF RUBRIC STRANDS FOR THE ASSESSMENT OF STUDENT WORK IN FIELD COURSE EXERCISES, WITH EXEMPLAR STATEMENTS

Task element	Knowledge, skills, and dispositions	Exemplar/mastery	Assessment structure
Lithologic description	Knowledge, skill	Description includes accurate information on rock type, mineralogy, grain size and range, texture, and special characteristics, in clear language with proper syntax and grammar.	Knowledge, lexical
Structural representation	Knowledge, skill	Structural interpretations are directly supported by measurements; inferred structures are distinguished from those directly observed; both small- and large-scale structures are represented.	Prototype
Symbology/ marking	Skill	Correct symbols and markings are used for structural features, contacts, internal features; these symbols show proper orientation and position; appropriate density to support inferences; clear and unambiguous representation of measurements and observations; measurements include all important features of base-map area.	Prototype
Presentation	Skill, disposition	Clean, neat; meets or approaches professional standards; layout of legend, key, etc., is clear and supportive of map presentation. Attention to detail is evident.	Prototype
Field book	Skill, knowledge, affect	Majority of both major and minor features are captured through complete written and graphical descriptions; measurements and observations are organized for easy review, retrieval, and interpretation; handwriting is clear and legible.	Knowledge, lexical
Supporting materials, e.g., stereonet plots, data tables, etc.	Skill, disposition	Supporting materials are directly tied to specific inferences; measurements (scale, angles, etc.) are accurate; materials are clear/focused and legible.	Knowledge, prototype

to learn from an activity, or at least what is expected of them through rubrics, the formative information can be better supplied, and the summative information will be more satisfactory for students and faculty alike.

4. Assessments require attention to outcomes, and to the experiences that lead to those outcomes—Assessments should encompass a full component of instructional planning and delivery, and never be far from the forefront for the group and the individual student, particularly when linked in a developmental sequence that serves long-range goals.

5. Assessments are valuable as both ongoing as well as episodic tools—Constant low-stakes formative assessments provide clarification to both students and instructors, and summative, episodic assessments signify completion of tasks.

6. Assessments should make a difference with issues of use and illuminate personal questions—With particular attention to inquiry skills and metacognitive abilities, assessment information can address such questions as “How do I do this?” “When am I going to use this?” and “How do I know when I am done?” Given that field time is often limited or costly, answers to these questions should be part of the set of dispositions for students.

7. Assessments should document and communicate successes, growth, and experiences to instructional and public audiences—To the extent that faculty use assessment data to improve future offerings of field course experiences, and document the success of program completers, a high value for the effort and resources committed can be demonstrated.

In applying our rubric to the field course experience tasks, it is our intention that these points are evident, which will contribute to students’ increased understanding of their tasks and the ways in which their learning was assessed. Attention to these points will also enhance the utility of the assessment data in the overall evaluation framework.

ASSESSMENT, INSTRUCTORS, AND INSTRUCTION

As previously stated, program evaluations that provide meaningful information collect data from a variety of sources and data that represent a variety of participants, faculty being one of these groups. It is generally expected for the design of field course experience activities to adhere to the goals of the course, but it seems a disservice to both the faculty and the program as a whole to limit faculty evaluation data to summative, end-of-course student evaluations of instruction. There are biases inherent in the administration and use of these instruments in higher education classrooms, as has been documented (Fox and Hackerman, 2003). However, if these instruments are biased, there is no guarantee that anecdotal information from student written comments is any less biased. Typically, these instruments are designed for in-class use and do not necessarily reflect the complexity of instruction in field course experiences, nor do they necessarily capture student responses relative to skills or dispositional learning. With the nearly full-time contact between faculty and students in field course experiences, there is the real prospect of an atmosphere in which personality is a contributor to recollection of past activities, by both students and faculty. If student assessments are to be explicit and largely objective, then faculty assessment as a function of evaluation should employ a more rigorous methodology that can demonstrate both validity and reliability.

As described already, the 2008 JMU field course experience was a transitional year, bringing in a variety of faculty new to both the geological as well as instructional context. Drawing on the faculty expertise, elements that were previously piloted, such as electronic data collection and mapping techniques, reached full implementation alongside traditional field mapping experiences. In addition, an environmental science–oriented module was piloted, based on reconnaissance during the previous summer. Coupled with the addition of four new faculty members on 2 wk rotations, a rather complex, and perhaps incomplete, set of interactions was imposed on both faculty and students. Add to this mix demands of driving field vehicles on the opposite side of the road, and opportunities for personalities to color both student and faculty expectations became evident. Anecdotes in student written comments suggested that issues of personal convenience colored the value of the learning experience by students. As a result, it became clear that a less biased data collection procedure needed to be adopted for future offerings.

How to Collect Data—Clinical Supervision

A useful framework to consider as a model for data collection and analysis was defined by Acheson and Gall (1997), termed the clinical supervision model. This approach is based primarily in precollege classroom instruction, but the techniques are readily adaptable to higher education settings, and the data collection and analysis methods are adaptable to different situations. In addition, the information that is produced is valuable for both formative purposes in the internal evaluation of learning experiences, but it is also useful for external summative purposes, relating first-hand observations of instruction that can be tied directly to explicit goals.

There are three phases to the clinical supervision cycle: (1) pre-observation, where the observer and observee meet before the instruction and discuss what is to be learned, the approaches that will be used, and any concerns or prior observations that may originate from either party; (2) observation, in which the data are collected through one or more techniques (discussed in more detail in the following); and (3) postobservation, in which there is joint reflection on the instruction, guided by the data that was collected. A summary of the information from each of these three phases has an immediate effect on subsequent instruction (the formative function), but it also documents for external audiences the intended result of the instruction, what happened during instruction, and how data were used to improve instruction and presumably student learning.

Data collection in the clinical supervision model can take on several forms: (1) selective verbatim techniques, in which portions of the dialogue between students and instructors are recorded faithfully, such as the questions that are asked or the types of instructions provided to the students; (2) map-based techniques, where a field mapping area (or portion) is used as the base, but the movements of instructors and students, their duration, and type of interaction are recorded, and (3) wide-lens techniques, which include videotaping and audiotaping, and standardized checklists of instructional behaviors.

These sources of data are primarily focused on the instructor, but the clinical supervision model does not preclude the use of student work. Indeed, field course experience's generate unique sets of artifacts produced by students, including maps, field notes, and sample collections. While these are used primarily for student assessment, when used in conjunction with the explicit and implicit goals of instruction, they become a valuable reflection tool in the postobservation domain. Examples of each of these data sources can be seen in Table 6.

Collecting data from each of these sources in a single session or set of sessions would not be easy, or even possible in a field setting. Neither would such data collection be appropriate, as the pre-observation discussion is designed to determine exactly which techniques or combination of techniques would best be employed, given the nature of the instructional activities, issues of concern, and overall program goals. The postobservation discussion is intended to determine the information to be gained from the collected data, and if the selection of techniques was in

TABLE 6. CLINICAL SUPERVISION DATA COLLECTION APPLIED TO FIELD COURSE SETTINGS

Selective verbatim	
Instructor structuring statements	"Your task is to map the lithologic units, contacts, and major structural features of the beach from Point A to Point B."
Student questions	"How can I tell a joint from a fault?"
Instructor feedback	"That grain might be plagioclase, but how could you tell it from orthoclase?"
Map based	
Student movement	On a base map, time indexed notations indicate the locations, dwell-times, and movement tracks of students.
Instructor movement	On a base map, time indexed notations indicate the locations, dwell-times, and movement tracks of the instructors relative to the students.
Wide lens	
Videotaping	Ideally, this would be a video camera set up in a remote location, but this is more suited to a classroom or laboratory setting.
Audiotaping	Students or instructors carry a tape recorder in field to either "talk out" actions while at outcrop, or capture dialogue between students and instructor.
Standardized student evaluation of instruction	Standardized forms with quantitative (usually Likert-scaled) items asking students to rate instructional quality, expectations, curricula, etc.
Artifacts	
Instructor generated	Instructions for mapping assignment; syllabi; reflections on exercises.
Student generated	Student maps, and lithologic descriptions in written form; photographs; field notes, relative to other data sources above.

fact appropriate. As student learning progresses and assignments become more demanding, so then should the data collected techniques be varied. Student artifacts become more complex, structuring statements become more specific yet limited in extent, and wider and more varied terrain is to be mapped.

Analysis of the data collected can be, in the narrowest manner, focused on specific questions that instructors might have on the progress or student response to an exercise. In the broader quest for reliability, however, the framework offered by Fox and Hackerman (2003) describes characteristics that can be used in pre- and postconferences, observation, and analysis. These characteristics include:

1. Knowledge of subject matter—Does the instructor demonstrate:
   - Mastery of the general content principles?
   - Sufficient breadth of knowledge within specific contexts?
   - Genuine interest in the content?
2. Skill, experience, and creativity with a range of pedagogies—Does the instructor:
   - Communicate clear expectations to students on assignments?
   - Recognize when students have difficulties?
   - Encourage discussion between students, and between students and instructors?
   - Persistently monitor student performance through formal and informal assessments, probes, interrogatives, etc.?
3. Understanding and use of appropriate assessment tasks—Does the instructor employ:
   - Assessments that are consistent with objectives and long range goals?
   - Persistent data collection on student performance during an activity?
   - Techniques to determine the extent of learning throughout the course?
4. Engagement in professional interactions beyond class—Does the instructor:
   - Contribute to ongoing intellectual development of the students, in and out of class?
   - Promote metacognitive and self-evaluative strategies in students?
   - Advise students that are having difficulty with learning the content and skills?
5. Communicating the results of reflections as a part of scholarly activity—Does the instructor:
   - Systematically share the results of the analysis, interpretation, and improvements with others in manuscripts, papers, and presentations?

There is a temptation to use all of these characteristics as a part of a checklist, in order to produce a unitary framework across instructors, field settings, or field course experiences. This decision should be approached with caution, as the application of these questions in the analysis of instructor data should also have the specific goals and objectives of both the particular activity and the field course experience as a whole in mind. When attempting to integrate instructor data with student data in the overall program evaluation, there should be a distinction (although not necessarily an exclusion) between the assessment of instruction versus assessment of instructors. A checklist is all too often used for the latter purpose only, and that may not provide the type of information that a field course experience needs to demonstrate efficacy to external audiences.

With respect to faculty observations in 2008, a limited amount of data was collected in a wide-lens observational manner, shared in an informal manner, and with only general goals in mind. Subsequent reflection among faculty, particularly when considering student written comments, urged the adoption of a more explicit means of defining and collecting data, to be used to improve instruction. For 2009, a small portion of each day was to be reserved for faculty to confer, focusing on a clinical supervision cycle for each faculty member, meeting beforehand and afterward, and using the location base map as a starting point, as each exercise involves multiple days on site. These efforts are to be linked to course goals and student performance in order for the overall evaluation framework to be justified. With the same format of faculty rotation, there is a greater depth of contextual experience that can be relied upon. Thus, faculty preparation will include preparation in the use of selective verbatim techniques, such as systematically recording student questions for short intervals, faculty structuring statements, and faculty feedback on specific map tasks. To the extent feasible, the use of small audio recording devices will be employed as a wide-lens technique, capturing dialogue between faculty and students. Map-based techniques will be also be employed by faculty members, tracking students across the field area. Finally, the range of student artifacts themselves (e.g., maps, cross sections, lithologic descriptions, etc.) will be compared to the assessment data described here for correspondence of goals, instruction, and assessment. Given the range of faculty expertise and rotations to and from the field sites, each faculty member will become at least familiar with each of these techniques, and it will be preferred for them to become well-versed in at least one of them, both in terms of data collection as well as analysis of those data.

CURRICULAR DESIGN ELEMENTS

In the larger context of the ways in which students learn science, Bransford and his colleagues (2000) suggested that learning in science is dependent on three factors: (1) identification of student preconceptions, (2) practicing science through inquiry, and (3) metacognition. A professional geologist needs a high level of skill in each of these domains in order to work effectively, either independently or as part of a team in the field. Student preconceptions, alternative conceptions, and misconceptions are deep-seated and related directly to past experiences and actions. Unfortunately, the literature on earth science misconceptions lags well behind the other sciences (see Duit, 2006) and is largely limited to material from precollege students. Libarkin and Anderson (2005) have examined the declarative and procedural knowledge of undergraduate students through the geosciences concept inventory (GCI), but this instrument was intended to be used in large introductory geology courses. An instructor should ask, "What are the preconceptions of students in field course experiences?" A reasonable assumption would be that, since they had presumably mastered the basic knowledge, skills, and dispositions in previous courses, preconceptions held by students would be supplanted by scientifically sound and representative ideas. However, there seems to be little data to support that assertion. Field course experience evaluation frameworks thus should use an analysis of student preconceptions to inform instructional design.

As was stated already, the goals and objectives of field course experiences are intended to be oriented toward the knowledge, skills, and dispositions necessary to geologists. Thus, the nature of geoscience inquiry is of high importance. When learning and practicing the use of equipment in the field, making and recording systematic observations, and making reasonable interpretations, students are engaged in the forms of inquiry that are conventional to the geosciences (Kitts, 1977; Frodeman, 2003; Pyle, 2008). Since the bulk of student objectives and assessment in field course experiences are skill-focused, it is appropriate that these data be used as a part of program evaluation, especially since assessment data may be cross-referenced to course goals and faculty actions.

Finally, the decision of the skills to employ, the knowledge to access, and persistence to a task are all driven by the executive, or metacognitive, function. Complete mastery is not a necessary prerequisite to field course experience tasks, but a student who has been prepared in a manner that integrates geoscience knowledge, skills, and dispositions, scaffolded from their preconceptions to strong geoscience metacognition, can begin to recognize the skills and knowledge to access in a given field situation. This function is often assumed to have occurred within successful students, and it may well be used as a part of program assessment when consulting alumni, but an analysis of the sense that students have of their increased knowledge, skills, and dispositions has largely been undocumented in the evaluation of field course experiences. Thus, if "learning" is to be documented as a part of an evaluation, it would be well served to include information on metacognition, particularly with respect to student skills and dispositions. If such an evaluation were to include clear documentation of changed student metacognition related to field course experience goals, any case for curricular decisions would be that much more compelling. A summary of these elements, cross-referenced with learning objective categories suggesting how work by Bransford and his colleagues (2000) can be applied to field course experiences, is presented in Table 7.

In prior offerings of the JMU field course experience, the precourse questionnaire asked students to indicate whether they had taken certain core courses or not. For the 2008 offering, this same information was collected, along with a request for their personal feelings of competence with the content represented in these courses, as a proxy for potential preconceptions. These data do not provide strong information on student preconceptions, but they do suggest that it would be fruitful to probe deeper into students' knowledge base, particularly in course areas (1) that they feel particularly comfortable with, (2) that they may be uncomfortable with, and (3) the intersection of these areas with field course experience objectives. These data will be collected from the 2009 field course experience participants. By sampling KSDs from among the KSDs inherent in the core courses, informal interviews with students will focus on preconceptions before a given exercise and on metacognitive strategies after an exercise.

The 2008 offering can be seen as a high-water mark between a traditional orientation toward analog geologic mapping skills, and one that is inclusive of both traditional as well as digital techniques. These were implemented as complementary techniques throughout the curriculum. During 2008, however, an environmental science strand was piloted in which each student participated, and whereby geologic mapping techniques were complemented by field techniques in stream and landslide geomorphology. This was based in part on perceived student interests and needs, and this was underscored by the data collected from the crude measurements employed at the onset of the field course experience and drove an evolution toward curricular change. As a result, the 2009 curriculum will develop in students a common set of traditional as well as digital mapping skills, and then allow them to select either a geologic or environmental science track that is geared more toward independent work. This design is built around a model that is intended to develop habits of mind as much as it is to solidify skills and enhance knowledge, embedding students first in a structured inquiry setting (Bell et al., 2005), and then into a more guided setting. As students progress toward the final weeks of the course, they will be engaged in independent mapping or environmental projects, where not only

TABLE 7. APPLICATION OF THE "HOW STUDENTS LEARN SCIENCE" FRAMEWORK TO STUDENT OBJECTIVE AND ASSESSMENT CATEGORIES

Domain	Preconceptions	Inquiry	Metacognition
Knowledge	Factual knowledge, use of terminology	Applies terminology to new situations in order to analyze situation or synthesize interpretations	Uses and adopts new terminology and concepts in novel situations
Skills	Use of compass, hand-lens, other tools	Designs and conducts investigation through a variety of data sources	Communicates with confidence the results of work in written and visual form
Dispositions	Ability to measure and record data and observations accurately and consistently	Consistently applies skills and knowledge with integrity, generates and tests multiple hypotheses and interpretations	Expresses clear self-evaluation of abilities, strengths, and weaknesses

will they be expected to produce detailed work on their own, but also defend and critique the work, promoting metacognition with respect to their own efforts.

DISCUSSION

In the context of evaluation, the combined impact of the analysis of student assessment, faculty clinical supervision, and attention to curricular design elements provides a triangulation of effort that establishes the "worth" or value of a given field course experience. A curricular design that is based on how people learn science can aid in the establishment of explicit, measureable knowledge and skill objectives, while at the same time providing at least indirect information on less explicit dispositions-based objectives. The objectives are then the "what" of the field course experience. When examining faculty actions relative to these objectives, the data become a clear basis for establishing the "how" of the field course experience. An analysis of student assessment data relative to the objectives, when combined with the analysis of faculty actions, contributes to understanding whether or not the "why" of the field course experience is met. Evaluation can document student success at meeting goals, identifying areas in need of improvement or development within the field course experience, and providing an analysis of cost-benefit ratios from the perspective of student performance, faculty resources, and instructional design.

Evaluation can be a time-consuming enterprise and perhaps seen as distracting from the main mission of instruction, but sound evaluation can also provide the basis for responses to key issues. Zimpher (1998) offered several key challenges that evaluation frameworks should be prepared to address. Applied to field course experiences, they become the basis for evaluation questions:

1. Teaching has and will receive more public scrutiny, and is more open to inspection than in the past. Student learning of the knowledge, skills, and dispositions provided by field course experiences should be the primary focus. In the context of a degree program, one should determine the extent to which field course experience goals contribute to programmatic goals. In addition, as students often seek field course experiences away from their home institution, attention should be paid to the extent to which these goals are recognized as valid by other degree programs.

Field course experiences are costly, both to the students as well as to the institution. They are resource-intensive on personnel, vehicles, and equipment. Through evaluation, one should ask if the field course experience is offered at an appropriate cost-benefit ratio.

2. Anecdotal reports are no longer sufficient by themselves because they are biased either by recollections or by selection of likely favorable anecdotes. As a function of even loose comparisons between field course experiences and other courses offered within a program, a single form of data or evidence is insufficient for comparison. Rather than compare apples with oranges, one should compare fruit baskets for sufficient sample comparison, particularly if students have the latitude to select from among a range of field course experience offerings.

Congruence between perceived employer expectations of professional knowledge, skills, and dispositions of graduates and the learning offered by the field course experience should also be documented, particularly when defining instructional goals.

3. Traditional assessments in addition to student and instructor artifacts are needed, depending on the range and specificity of field course experience goals. Quantitative measures provide information on gains relative to specific content, but when complex interactions of knowledge, skills, and dispositions are the goal, and professional self-awareness is an outcome, more types of data are needed to triangulate toward assertions of quality or efficacy.

4. Content transmission will be less of the focus. If a field course experience is to be a capstone or synthesis experience, the focus shifts from basic content transmission to helping learners access information and collect basic data and observations necessary to the context of investigation.

Preconceptions held by students should be determined, so that they do not impede development of skills and dispositions. Student-constructed solutions should be directed toward self-evaluation strategies that will develop metacognitive strategies.

5. Curriculum design should be linked to teaching and learning. Linking teaching and learning requires coordination of goals, content, and teaching, such that faculty work together in the articulation of goals and objectives within an overall program.

To facilitate such learning, the instructional team must share a clear understanding of the curricular elements that best promote student learning in order to provide the instruction that supports this learning.

6. Students have experienced a range of pedagogies. Prior to the field course experience, students have experienced a range of pedagogical approaches, from teacher-centered lecture and guided laboratory experiences to field settings. The experiential nature of field learning should provide a broad range of experiences matched to the expected knowledge, skills, and dispositions.

Students have changed expectations about the nature of quality teaching, and because of their varied experiences, they need to see how experiences are tied to the goals of the field course experience.

*7. There is a new scholarship of teaching and learning.*Where high value is placed on the scholarship of teaching, faculty must systematically pose questions of their teaching, selecting the means and methods of collecting the data, and analyze the data in an appropriate manner (Boyer, 1990). To the extent that models of teaching and learning in a field course experience are well documented, faculty should communicate the results of their research to other practitioners, to apply and or to replicate the results.

In meeting these increased and broadened expectations for a field course experience, it would be useful then to use the elements discussed herein as a sort of "tool kit" for evaluating and influencing the development of a field course experience that can meet the challenges stated previously. First, objectives in each domain (knowledge, skills, and dispositions) should be constructed in a measureable manner and closely linked to the assessment criteria associated with each task in the field course experience. These objectives and assessment criteria are not only made explicit to students, but they are also compared by faculty with student preconceptions, so that appropriate instructional decisions may be made.

Second, faculty should be provided with a data-oriented framework to reflect upon when considering each field exercise, examining how their actions help students to meet the stated objectives in a mindful and efficient manner. Data collected as a part of a clinical supervision cycle should not be viewed as evaluative in and of themselves. Rather, they should be viewed as an additional, unbiased data set, aimed at the learning goals of the field course experience, and assisting in the continual development of the field course experience curriculum. To the extent that faculty work together in collecting these data, their individual expectations can be made explicit, and a clear team approach to instruction can be realized.

Finally, a mindful approach to the design of a field course experience curriculum provides the means by which the support for student learning progress through a field course experience can be clearly documented. To the extent that the design is guided by contemporary research on how students learn science, it is more likely that students will meet the intended learning goals. When student preconceptions are considered as an element of instructional design, the nature of the scientific inquiries that are made available to students by faculty may be tailored in such a way that student metacognition is the result and the professional mindset sought as a result of a field course experience is realized. Documentation of this process in the development and evaluation of curricular materials is of demonstrable value in achieving learning goals (Kesidou and Roseman, 2002).

One consideration in comprehensive evaluation frameworks has become increasingly important in the last few years, especially where data on human participants is to be included. Each institution in the United States where research activities are conducted is expected to have an Institutional Review Board, which oversees and approves research conducted with human participants. If the evaluation plan is implemented for purely internal reasons, at either the program or institutional level, then it is normally not considered "research." However, the drive for faculty to document a scholarship of teaching makes evaluation information valuable to a broader professional audience, and this transforms an evaluation project into generalizable research. This then requires faculty to be trained to recognize the rights of those participants, by obtaining from students their informed consent for the information to be used for research. Sanctions for noncompliance can be severe, including an institutional requirement for publications using data obtained without consent to be retracted. Each institution that receives federal funding is subject to these regulations.

CONCLUSIONS

The need to develop and employ an evaluation framework in educational programs is a necessity for both internal curricular decisions as well as external documentation to administrators. Student assessments of learning are a feature of any course, and the nature of field course experiences demands a unique format for assessment that includes not just student knowledge, but also a clear documentation of their growth in scientific skills and professional dispositions. Each of these factors is fundamental to a field course experience, and assessments that lack skills and dispositional aspects are incomplete. Assessments should attend to the literature on the methods with which individuals learn science, starting with their preconceptions and ending with their metacognitive skills, and do so as a normal part of instruction. Assessments should also take into account the complex verbal and visual nature of field course experiences, being based on clear and explicit expectations transmitted to students.

The role of faculty relative to the curriculum, the students, and the exercises on-site is seldom examined in the context of field course experiences, but it is included in a growing field in higher education science instruction in general. At the same time, the limitations of traditional student evaluations of instruction have been realized, making the need for rigorous, alternative forms of collecting data for formative and summative purposes much more evident. In addition to this situation, the nontraditional context of field course experiences and the difficulties of producing these data only increase.

Together, both student and faculty data are necessary for an effective evaluation; once a curriculum is established and delivered, the match of student performance and learning relative to the intentions of the faculty must be determined. The relationship of field course experience learning experiences to overall undergraduate program goals and the expectations of the profession should be continually demonstrated in order to justify assertions of professional value for a field course experience to those that hold the purse strings. A comprehensive evaluation plan, designed and implemented by those who are responsible for the field course experience, is a means to accomplish this, providing a richer data set for continuing improvement and adjustment than a generic evaluation template, generated for more traditional instructional models.

APPENDIX 1. SELECTED FIELD COURSE EXPERIENCE OBJECTIVES AND OUTCOMES, ACCESSED THROUGH KING (2009).

Illinois State University, Northern Illinois University, and Western Kentucky University

1. To learn basic field techniques, particularly: using the Brunton compass, measuring geologic sections, describing rocks, taking field notes, and making field sketches.
2. To learn the latest technologies that are used in the construction of geologic maps. Participants will be introduced to using PDAs [personal data assistants] equipped with blue-tooth GPS units to gather and analyze field data.
3. To learn the skill of geologic mapping, a process that involves total immersion in the science and in the project at hand, and the associated skills of location on topographic maps and air photos and interpretation of features.

4. To learn to interpret the structure and geologic history of an area based on field observations and geologic map. Such ability is demonstrated mainly through the construction of geologic cross sections from geologic maps.

5. To learn the importance of accuracy in data acquisition and placement on a geologic map.

6. To integrate aspects of prior coursework into a comprehensive package in which the student becomes aware of the interdependence of all parts of the science of geology.

7. To develop an appreciation of the scale of geologic features and of the "reality" of geologic features, as compared to their depiction in print media.

8. To develop the skills and expertise needed to make the transition from student to professional geologist.

9. To develop senses of self-confidence and professional competence.

Lehigh University

The goal is to provide a synoptic, capstone field experience for geology and environmental science majors, and instruction on how to make, read, and interpret geologic maps and how to envision field problems and collect environmentally diagnostic data. The field, field geologic relationships, and the concepts of geological mapping and environmental data are used as the vehicle toward development of a professional earth and environmental scientist.

Georgia State University

1. To see illustrated the classic theoretical concepts of geology.

2. To learn the basic field skills necessary for any field study in earth/environmental sciences.

3. By actually making a map, to learn techniques of how to read and gain the maximum amount of information from published maps.

James Madison University

After completing the field course, students will be qualified to work for an industrial, governmental, or academic employer who needs individuals to make their own way to an isolated village in a foreign country, assess the local geology, natural resources, natural hazards, environmental conditions, etc., write a project report, draft a publishable map, generate a data base, and return home safely. The main objective is for the participant to become confident at scientific observation, interpretation, and solution of geological problems in the field. Participants will learn to recognize and interpret a wide variety of rock types, structures, and geomorphic features. Emphasis is placed on methods of map-making, data recording, and report preparation. Projects from one to five days duration will be conducted in well-exposed igneous, metamorphic, and sedimentary rocks, ranging in age from Precambrian through Quaternary, and correlative rocks and sediments of the northern Appalachians.

Bowling Green State University

The course will teach students how GPS navigation and digital mapping and data analysis using geographic information systems (GIS) can facilitate fieldwork and improve the understanding of the geology. Working with sedimentary, metamorphic, and igneous rocks, students learn how to make methodical observations, accurate recordings, and sound interpretations of the geology seen in outcrop. Exercises include measurement and analysis of sedimentary sections, construction of geologic maps, structural analysis of folds and faults, slope stability analysis, and environmental assessments. Students will learn to use Brunton compasses, laptop, and ruggedized tablet PC computers (Xplore Technologies), GPS receivers, aerial photographs, topographic maps, satellite images, and GIS databases in their projects. Field areas are in the Basin and Range, Colorado Plateau, and Rocky Mountain provinces. Geologic features to be examined are folded and faulted sedimentary strata of Paleozoic and Mesozoic age, regional metamorphic facies in Precambrian rocks, volcanic domes and pyroclastic rocks of Tertiary age, pegmatites and plutonic rocks of Precambrian age, and Quaternary glacial deposits. Environmentally related projects include slope stability analysis and environmental site assessments.

Michigan Technological University

This study abroad program to East Africa is intended to serve several purposes: (1) give student a hands-on knowledge of the geology and geological processes in the East African Rift Valley, (2) provide an alternative for geology students needing a geology field camp, and (3) help the curious understand and appreciate one of the geologic marvels of our time, the East Africa Rift Valley.

West Virginia University

1. To learn how to describe and log stratigraphic sequences of sedimentary rocks.

2. To learn how to construct a geologic map of an area comprising several square kilometers. Students use topographic base maps, aerial photos, GPS units, and compasses to map two separate areas encompassing a variety of folded and faulted sedimentary rocks as well as igneous intrusions.

3. Additional goals include: gaining confidence in making geologic observations and interpretations; broadening geologic experience beyond the classroom; and learning to deal with incomplete or missing data.

4. Geology 404 is a capstone experience that requires students to demonstrate mastery of the concepts and skills acquired during the undergraduate years.

University of Hawaii

1. Students can explain the relevance of geology and geophysics to human needs, including those appropriate to Hawaii, and are able to discuss issues related to geology and its impact on society and planet Earth.

2. Students can apply technical knowledge of relevant computer applications, laboratory methods, and field methods to solve real-world problems in geology and geophysics.

3. Students use the scientific method to define, critically analyze, and solve a problem in earth science.

4. Students can reconstruct, clearly and ethically, geological knowledge in both oral presentations and written reports.

5. Students can evaluate, interpret, and summarize the basic principles of geology and geophysics, including the fundamental tenets of the subdisciplines, and their context in relationship to other core sciences, to explain complex phenomena in geology and geophysics.

REFERENCES CITED

Acheson, K.A., and Gall, M.D., 1997, Techniques in the Clinical Supervision of Teachers (4th ed.): New York, Longman, 288 p.

Anderson, C.W., 1987, Strategic teaching in science, *in* Jones, B.F., et al., eds., Strategic Teaching and Learning: Cognitive Instruction in the Content Areas: Alexandria, Virginia, Association for Supervision and Curriculum Development, p. 73–91.

Anderson, D.S., and Miskimins, J.L., 2006, Using field-camp experiences to develop a multidisciplinary foundation for petroleum engineering students: Journal of Geoscience Education, v. 54, no. 2, p. 172–178.

Astin, A.W., Banta, T.W., Cross, K.P., El-Khawas, E., Ewell, P.T., Hutchings, P., Marchese, T.J., McClenney, M., Mentkowski, M., Miller, M.A., Moran, E.T., and Wright, B.D., 1996, Assessment forum: 9 principles of good practice for assessing student learning: Washington, D.C., American Association for Higher Education, http://www.facet.iupui.edu/resources/AAHE%20Principles.pdf (accessed 11 March 2009).

Atkin, J.M., and Coffey, J.E., 2003, Everyday Assessment in the Science Classroom: Arlington, Virginia, National Science Teachers Association, 184 p.

Ault, C.R., Jr., 1998, Criteria of excellence for geological inquiry: The necessity of ambiguity: Journal of Research in Science Teaching, v. 35, no. 2, p. 189–212, doi: 10.1002/(SICI)1098-2736(199802)35:2<189::AID-TEA8>3.0.CO;2-O.

Ajzen, I., and Fishbein, M., 1980, Understanding Attitudes and Predicting Social Behavior: Englewood Cliffs, New Jersey, Prentice-Hall, 278 p.

Baker, M.A., 2006, Status Report on Geoscience Summer Field Camps: Washington, D.C., American Geological Institute, Report GW-06-003, 8 p.

Bell, R.L., Smetana, L., and Binns, I., 2005, Simplifying inquiry instruction: Assessing the inquiry level of classroom activities: Science Teacher (Normal, Illinois), v. 72, no. 7, p. 30–33.

Bloom, B.S., 1956, Taxonomy of Educational Objectives. Handbook I: The Cognitive Domain: New York, David McKay Co., 196 p.

Boyer, E.L., 1990, Scholarship Reconsidered: Priorities of the Professoriate: Princeton, New Jersey, Carnegie Foundation for the Advancement of Teaching, 160 p.

Boyle, A., Maguire, S., Martin, A., Milsom, C., Nash, R., Rawlinson, S., Turner, A., Wurthmann, S., and Conchie, S., 2007, Fieldwork is good: The student perspective and the affective domain: Journal of Geography in Higher Education, v. 31, no. 2, p. 299–317, doi: 10.1080/03098260601063628.

Bransford, J.D., Brown, A.L., and Cocking, R.R., eds., 2000, How People Learn: Brain, Mind, Experience, and School: Washington, D.C., National Academies Press, 374 p.

Butler, R., 2008, Teaching geoscience through fieldwork: Bradford, UK, Geography, Earth, and Environmental Sciences (GEES), http://www.gees.ac.uk/pubs/ guides/fw/fwgeosci.pdf (accessed 11 March 2009).

Bybee, R.W., ed., 2002, Learning Science and the Science of Learning: Arlington, Virginia, National Science Teachers Association, 151 p.

Chamberlin, T.C., 1897, The method of multiple working hypotheses: The Journal of Geology, v. 5, p. 837–848.

Chi, M.T.H., Feltovich, P.J., and Glaser, R., 1981, Categorization and representation of physics problems by experts and novices: Cognitive Science, v. 5, p. 121–152.

Chiappetta, E.L., and Koballa, T.R., 2006, Science Instruction in the Middle and Secondary Schools: Developing Fundamental Knowledge and Skills for Teaching (6th ed.): New York, Allyn and Bacon, 320 p.

Dave, R.H., 1975, Developing and writing behavioral objectives (Armstrong, R.J., ed.): Tucson, Arizona, Educational Innovators Press.

Duit, R., 2006, Students' and Teachers' Conceptions and Science Education: Kiel, Germany, Institut für die Pädagogik der Naturwissenschaften, http://www.ipn.uni-kiel.de/aktuell/stcse/stcse.html (accessed 11 March 2009).

Ebert-May, D., 1998, Annotated Version of Potential Activities and Assessment for Faculty Professional Development Workshops: http://darkwing.uoregon.edu/~first/examples.htm (accessed 5 September 2009).

Fox, M.A., and Hackerman, N., eds., 2003, Evaluating and Improving Undergraduate Teaching in Science, Technology, Engineering, and Mathematics: Washington, D.C., National Academies Press, 232 p.

Frechtling, J., 2002, The 2002 User-Friendly Handbook to Project Evaluation: Arlington, Virginia, National Science Foundation, Report NSF 02-057, 92 p.

Frodeman, R., 1995, Geological reasoning: Geology as an interpretive and historical science: Geological Society of America Bulletin, v. 107, no. 8, p. 960–968, doi: 10.1130/0016-7606(1995)107<0960:GRGAAI>2.3.CO;2.

Frodeman, R., 2003, Geo-logic: Breaking the Ground between Philosophy and the Earth Sciences: Albany, New York, State University of New York Press, 192 p.

Gardner, H., 1993, The Unschooled Mind: How Children Think and How Schools Should Teach: Cambridge, Massachusetts, Basic Books, 320 p.

Goetz, J.P., and LeCompte, M.D., 1984, Ethnography and Qualitative Design in Educational Research: San Diego, California, Academic Press, 292 p.

Hemler, D., and Repine, T., 2006, Teachers doing science: An authentic geology research experience for teachers: Journal of Geoscience Education, v. 54, no. 2, p. 93–102.

Hughes, P., and Boyle, A., 2005, Assessment in the Earth Sciences, Environmental Sciences, and Environmental Studies: Bradford, UK, Geography, Earth, and Environmental Sciences (GEES), 47 p.

Huntoon, J.E., Bluth, G.J.S., and Kennedy, W.A., 2001, Measuring the effects of a research-based field experience on undergraduates and K–12 teachers: Journal of Geoscience Education, v. 49, no. 3, p. 235–248.

Kesidou, S., and Roseman, J.E., 2002, How well do middle school science programs measure up? Findings from Project 2061's curriculum review: Journal of Research in Science Teaching, v. 39, no. 6, p. 522–549, doi: 10.1002/tea.10035.

King, H.M., 2009, Geology field camps—Comprehensive listing: http://geology.com/field-camp.shtml (accessed 10 March 2009).

Kitts, D.B., 1977, The Structure of Geology: Dallas, Texas, SMU Press, 199 p.

Kizlik, B., 2009, Measurement, Assessment, and Evaluation in Education: http://www.adprima.com/measurement.htm (accessed 10 March 2009).

Krathwohl, D.R., Bloom, B.S., and Masia, B.B., 1973, Taxonomy of Education Objectives, the Classification of Educational Goals. Handbook II: Affective Domain: New York, David McKay Co., 196 p.

Lawrence, S., and Margolis, E., 1999, Concepts and cognitive science, *in* Margolis, E., and Lawrence, S., eds., Concepts: Cambridge, Massachusetts, Massachusetts Institute of Technology Press, p. 3–83.

Libarkin, J.C., and Anderson, S.W., 2005, Assessment of learning in entry-level geoscience courses: Results from the geoscience concept inventory: Journal of Geoscience Education, v. 53, no. 3, p. 394–401.

Mestre, J.P., and Cocking, R.R., 2000, The science of learning: Journal of Applied Developmental Psychology, v. 21, no. 1, Special Issue, p. 1–135.

Murphy, G.L., 2002, The Big Book of Concepts: Cambridge, Massachusetts, Massachusetts Institute of Technology Press, 563 p.

Noll, M.R., 2003, Building bridges between field and laboratory studies in an undergraduate groundwater course: Journal of Geoscience Education, v. 51, no. 2, p. 231–236.

O'Neal, M.L., 2003, Field-based research experience in earth science teacher education: Journal of Geoscience Education, v. 51, no. 1, p. 64–70.

Pyle, E.J., 2008, A model of inquiry for teaching earth science: Electronic Journal of Science Education, v. 12, no. 2, http://ejse.southwestern.edu/volumes/v12n2/articles/art1-pyle.pdf (accessed 11 March 2009).

Resnick, L.B., 1983, Mathematics and science learning: A new conception: Science, v. 220, p. 477–478, doi: 10.1126/science.220.4596.477.

Sibley, D.F., 2005, Visual abilities and misconceptions about plate tectonics: Journal of Geoscience Education, v. 53, no. 4, p. 471–477.

Simpson, E.J., 1972, The Classification of Educational Objectives in the Psychomotor Domain: Washington, D.C., Gryphon House.

Smith, E., 1995, Concepts and categorization, *in* Smith, E., and Osherson, D., eds., Thinking: An Invitation to Cognitive Science, Volume 3: Cambridge, Massachusetts, Massachusetts Institute of Technology Press p. 3–33.

Trowbridge, L.W., Bybee, R.W., and Powell, J.C., 2004, Teaching Secondary School Science: Strategies for Developing Scientific Literacy (8th ed.): Upper Saddle River, New Jersey, Prentice Hall, 384 p.

Zimpher, N., 1998, Ten changing demands on college teachers in the future, *in* Changing Demands on College Teachers: A Conference on Teaching Support Providers, 27 April, Columbus, Ohio: http://web.archive.org/web/20040606130612/http://www.acs.ohio-state.edu/education/ftad/Publications/ten-nancy.html.

Manuscript Accepted by the Society 5 May 2009