From **Gal, I. & Garfield, J. B. (editors). The Assessment Challenge in Statistics Education**. IOS Press, 1997 (on behalf of the ISI). Pages 165-178. ISBN 90 5199 333 1.  The References section is a separate chapter and is available on the Internet at http://www.stat.auckland.ac.nz/~iase/publications/assessbkref.

# 13. Portfolio Assessment in Graduate Level Statistics Courses

**Carolyn M. Keeler**

## Purpose

In this chapter, the process of developing a form of alternative assessment, the portfolio, will be described. The portfolio is a purposeful collection of student work that exhibits the student's efforts, progress and achievements over time. Portfolio development supports the assessment of long-term projects, encourages student-initiated revision and provides a context for presentation, guidance, and critique. The purpose of portfolio development is the same no matter the course or age of the students, to display the products of instruction in a way which challenges teachers and students to focus on meaningful outcomes. The context in which the use of portfolios is described here is a graduate level statistics course where an additional purpose is to provide students with an organized reference on statistical programming, analysis, and interpretation. However, the process used in developing portfolios and the important issues surrounding portfolio assessment can easily be generalized to different educational levels and subject areas. Some of the questions addressed in this chapter are 1) What is the underlying belief concerning knowledge construction which guides portfolio assessment? 2) How do you develop and use portfolios? 3) What does a portfolio look like? and 4) What are the major considerations in deciding to use portfolio assessment?

## INTRODUCTION

Teachers are beginning to recognize that traditional tests are limited by easy scoring formats and the testing of easily retrievable fragments. The skills that result from students passively absorbing information and taking multiple choice or true-false tests may have little or no relevance outside of the classroom because they are not transferred to everyday problems. Research from the field of psychology indicates that learning does not occur by passive absorption alone. In many situations, people learn using prior knowledge, assimilating new information, and constructing their own meanings (Resnick, 1987). This constructive, active view of the learning process is reflected in the way much of mathematics is taught. In rewriting curriculum to allow for the construction of meaning by the learner and the application of learning

to solve real-life problems, the dilemma is that we must also change the way we assess curricular obtainment. Assessment has the potential to teach and gives the student opportunities to develop complex understandings.

Good assessment enables us to accurately characterize students' functioning and performance in order to make sound decisions to improve instruction. Portfolio assessment is an example of assessment integrated with instruction and is one alternative to one-dimensional testing. The Encyclopedia of Educational Evaluation (Anderson, Ball, & Murphy, 1975) defines assessment as a process built around multiple indicators of performance:

> Assessment, as opposed to simple one-dimensional measurement, is frequently described as multitrait-multimethod; that is, it focuses upon a number of variables judged to be important and utilizes a number of techniques to assay them...Its techniques may also be multisource...and/or multijudge (p. 27).

Portfolios are by their nature multisource. Demonstration of what students know and are able to do takes many different forms in the portfolio. Students, teachers, or students and teachers together may choose what example to include of the activities completed for each different task. Although some tasks may have been handed in previously and revised, the student writes the reasons for including the example and a critique of what was learned and demonstrated (or not). This is both a learning and assessment strategy: students learn through the critique of their work and demonstrate the level of understanding they now have of the concept. This in turn informs the teacher of the need for further clarification or reteaching.

With some variation, the process for developing portfolio assessment is the same as that used by developers of performance measurements no matter what their nature. (See also chapter 3 in this regard.) This process is offered here in seven steps. It is generally accepted that the process should conclude with the revision of both instructional and assessment tasks.

Step 1 Delineate the knowledge and performance and process skills students are to develop.
Step 2 Develop assessment tasks which require students to demonstrate the knowledge and skills delineated.
Step 3 Specify the criteria for judging student performance on each task.
Step 4 Seek evidence of validity in measuring the knowledge and skills delineated in the course content through expert panel or item analysis.
Step 5 Develop a reliable rating process or rubric for each task.
Step 6 Provide feedback to students.
Step 7 Refine assessment and improve curriculum and instructional tasks.

The chapter is organized under four headings, 1) the constructivist framework, 2) portfolio development, 3) portfolio sketch, and 4) implications, issues and concerns. The section on the constructivist framework describes this theoretical basis for instruction. Students are asked to construct their own understanding of the subject matter through active involvement in the learning environment using authentic learning tasks and application of the knowledge gained to real-life predictable and unpredictable problems through outside projects. The portfolio development section includes the purpose of portfolio assessment, development, creating and critiquing assessment tasks, and establishing the scoring criteria.

In the portfolio sketch, one example of an application of the development process is offered. In the example, portfolio assessment is used to demonstrate mastery of computer software programming and the application of statistical processes to projects and exams. A reflective journal is included to emphasize the importance of student reflection on the learning strategies used to acquire the knowledge and application of knowledge to course products. Finally, the scoring procedure is described.

The final section covers implications for use of the portfolio and some issues and concerns which emerge as a result.

## THE CONSTRUCTIVIST FRAMEWORK

In the early 60' s following the launch of Sputnik in 1957, the focus of education in the United States became less on liberal arts and more on the scientific and theoretical base of knowledge. As a result, a philosophy called objectivism pervaded the methodology of teachers. It holds that knowledge exists independently of the mind for transfer into a student as if into an empty vessel. In the field of statistics this would be realized in teaching emphasizing the rote memorization of formulas. In the 90' s, the constructivist framework for education is becoming quite different. Rather than transmitting knowledge through texts and lectures and asking students to memorize facts, the model for the teaching-learning act becomes one of transaction, where meaning is made through a significant transaction with the information (Pulaski, 1980; von Glasersfeld, 1981; Davis, Maher, & Noddings; 1990; Schifter & Fosnot, 1992). The main idea is that learners must construct their own understandings rather than passively absorbing or copying the understandings of the instructor. In my practice, this constructivist view of learning is not only the theoretical basis for instruction but is also a principal goal of instruction. When learning is an active process the student demonstrates understanding through such products as reports, speeches, and models. When these products become part of a portfolio, they may be extended, revised, or reflected upon by the learner and thereby become one form of assessment used to determine the student's progress.

As students actively engage in problem situations, they build understandings which are an extension of and then become a part of current knowledge. Learning is the result of an action in which students participate, not the inevitable product of encountering materials. As part of this meaning-making, opportunities are provided for students to pause and reflect upon what they have been doing. The instructor can assist students through simple devices such as pausing and asking students to recount to a neighbor the problem solving process in which they have engaged, leading group discussions of the thought process of individual students, or having students reflect in a journal regarding the development of their own understanding.

The radical constructivist perspective suggests that the individuals' understandings are built from their unique web of prior concepts and their current subjective experience and therefore are idiosyncratic (Schifter & Simon, 1992). It is imperative then that the assessment of this unique understanding include problem solving opportunities which parallel learning activities and opportunities for reflective expression describing the individuals' changing conceptions. For this reason, many of the projects assigned as coursework are extended and reflected upon as part of the portfolio in the form of a reflective journal.

## PORTFOLIO DEVELOPMENT

## The purpose of portfolio assessment

Portfolios are purposeful collections of student work that are reviewed against criteria in order to judge an individual student or a program, or both. The products on display in the portfolio focus instruction for both the teacher and the student on critical learnings. The portfolio of itself does not constitute the assessment. The "assessment" in portfolio assessment only exists when three important things are communicated to the students, 1) the assessment purpose, 2) the criteria or methods for determining what is put into the portfolio, and 3) the evaluation criteria for scoring the pieces or the collection. A portfolio in statistics education may include statistical tests and interpretation of results, projects and exams which have been corrected, and a reflective journal (see Portfolio Sketch section later in this chapter).

The portfolio has a meta-structure allowing for both traditional and alternative sources of data or multiple indicators of the same outcomes and therefore offers a more complete picture of achievement. Because it is longitudinal, it has the advantage of putting one assessment or project into perspective and documenting growth over time. It also documents what concepts and models have been acquired through the course. Portfolios should include pieces representing student progress, student reflection about their work, and evaluation criteria (Herman, Aschbacher, & Winters, 1992). For evaluation purposes, the portfolio can do what traditional assessment in most cases does not do, provide direct evidence for evaluating the student's progress over time, at mastering the essential concepts and techniques of the course. Many methods of assessment included in other chapters can yield usable results; some of these are concept mapping (Chapter 8), group work (Chapter 10), performance assessments (Chapters 3, 11, and 12), and model-building (Chapter 6).

## Development

The process of portfolio development is dependent on the identification of the most important goals of instruction. In a graduate statistics course, both research models and statistical procedures are taught. Thus, students should be able to 1) choose the appropriate model and 2) formulate the research question in a way that allows them to 3) appropriately choose sampling and data collection procedures. Students should 4) know the appropriate statistical test to apply to the data and 5) how to calculate the test statistic and/or use a statistical package to computerize the analysis. Students should be able to 6) interpret the statistical output from the computer program and 7) draw conclusions from the results. (See chapter 1 for further discussion of common instructional goals in statistics education.)

Once you know what you want the students to know and be able to do at the completion of a course of study, the next step is to develop learning tasks which involve them in the discovery and application of this knowledge. Statistical solutions are not only a function of knowledge but also of metacognitive skills. It is important to allow time for self-monitoring, self-critique, and self-regulatory processes. Emphasis should be placed on making connections, understanding concepts, learning to reason statistically and developing the ability to communicate statistics effectively. The development of activities and projects which facilitate the student's ability to create personal meaning from new information and discover key concepts is the goal. Learning isn' t necessarily a linear progression of discrete skills.

Authentic activities and problems to solve must be developed by which students will 1) represent problem situations verbally and symbolically, 2) integrate statistical concepts and procedures, 3) apply strategies to solve problems, 4) formulate questions and analyze problem situations, and 5) verify and interpret results. Through their discovery of these processes they begin to internalize the skills and knowledge which make the next activity easier. You cannot assess performance unless you teach performance.

Once goals and learning processes are articulated, appropriate assessment tasks may be developed to match the active, inquiry-based learning process. Cognitive learning theory encourages linking instruction with assessment yet tells us that there is great variety in learning styles, approaches to problem solving, memory, developmental paces and aptitudes. Therefore, choices should be provided in task assignments and there should be ample time for project development, revision, and rethinking (Herman, Aschbacher, & Winters, 1992). Assessment tasks should target instructional aims and allow students to demonstrate their understanding and reflect on their progress. Just as engaging students in real-world tasks is motivating, it aids in the transferability of skills and knowledge.

Alignment between the curriculum and the assessment occurs in two ways, content and context. Content alignment is necessary for assessment validity. Context alignment occurs when the assessment protocol or scenario is the same as the one used to teach the curriculum (English, 1992). Context alignment is particularly important in statistics education. When students are given traditional tests in statistics there is a tendency for recall level knowledge and memorization of formulas to dominate learning rather than the application of statistical procedures to real-world problems. Portfolio assessment which is aligned with authentic learning tasks in class can be developed which help to transfer and extend knowledge by reinforcing the application of statistical procedures and the interpretation of results.

When students solve problems, make decisions, construct arguments, defend their answers, or collaborate with others, they are demonstrating the processes involved in the discipline while gaining understanding and developing a product similar to a professional in the field (Resnick & Klopfer, 1989). Modern curriculum theorists believe that engaging students in the processes of the discipline is a powerful learning strategy (Baker, Freeman, & Clayton, 1991; Bransford & Vye, 1989; Resnick & Klopfer, 1989). For example, the Content Assessment Prototype in history, developed by Baker and colleagues (1992) at CRESST, engages students in the authentic tasks of historians. In our case, students develop a “working knowledge”, one which enables them to use the tools for making, using, and communicating information through research.

Three more factors must be considered in designing a portfolio, the purpose, context and design. The purpose affects the content of the portfolio and so must be decided first. This includes what you want to show with the portfolio, i.e., mastery of concepts, understanding of the processes by which knowledge is constructed, attitudes towards statistics. Purpose also includes how the portfolio will be used, i.e., accountability, student self-analysis, program evaluation. Context includes such things as students’ characteristics, activities which occur during the time the portfolio is in production and the uses of the completed portfolio. Design covers such considerations as what will count as evidence, how much evidence is needed, how it will be presented, who will decide what evidence to include, and the evaluation or scoring criteria. What is included also depends on who will view the completed portfolio. Known as stakeholders, the teacher, student, and in the context of K-12 education, parents and administrators, may be included as audiences of the portfolio.

## Creating and critiquing assessment tasks

Unique to the creation of the portfolio as a form of alternative assessment, is the involvement of students in deciding what to include. The students should know "up front" how the portfolio will be weighted in the final grade and what assessment tasks, course projects, and exams are assigned as a part of the portfolio. Teacher and student should collaborate to decide what other things to include as representative of their learning and performance. Arter and Spandel (1992) summarize the kinds of questions teachers should keep in mind when creating portfolios:

- Is the work included in the portfolio representative of what students can really do?
- Do the portfolio pieces represent coached work? Independent work? Group work? Is the amount of support the student received indicated?
- Do the evaluation criteria for each piece and the portfolio as a whole represent the most relevant or useful dimensions of student work?
- Is there alignment between the course goals, the learning tasks and the assessment tasks? For example, does the task require multi-faceted cognitive skills if this is a course goal?
- Do tasks or some parts of them require extraneous abilities?

Before you use assessment tasks with students, it is a good idea to allow a period of time to lapse following development and then review the task to see if it is sound. Another review of the task should be made after initial use and scoring. Often problems with a task surface during use which were impossible to anticipate in development. The determination of the validity of the task and the reliability of the scoring rubric is simplified if you are the instructor of the course and all of the tasks will be scored by you. The teacher's knowledge of what the instructional objective is and therefore what should be assessed guarantees some degree of validity. If you are the only one using the scoring rubric, then you must be careful to be consistent between portfolios but inter-rater reliability becomes irrelevant. If you are developing tasks and scoring rubrics to be used by more than one instructor, the task should be written collaboratively and the first time the task is assigned both people should score the task and compare ratings. Some questions by which to review the tasks you have developed include:

- Can tasks provide reliable measurements? For example, is there inter-rater agreement and between-task generalizability?
- Are they sensitive to students? For example, do tasks, scoring processes, etc. avoid potential bias?
- Do all students have access to the opportunities and instruction which will prepare them to perform well?
- Are the tasks clearly delineated?
- Do they provide achievement information that differs from traditional measures? For example, does the assessment engage students in authentic or real-world tasks which are worth doing?
- Do the tasks offer different ways of assessing the same concept development or provide choices to students of tasks to complete?

## Establishing the scoring criteria

Assessments should be designed to improve performance not just measure it. Feedback to students must provide direct, usable information on their performance and concrete examples of the difference between their current performance and the expected performance. Where a single score on a test is unuseful, coded information from the teacher, a scoring criteria or rubric provides detailed feedback of performance in the analytic areas.

Students should know in advance the criteria for judging student performance. The rubric includes a description of the dimensions on which tasks will be scored, the scale which will be used in rating performance, and a description of the characteristics of each level of the scale. The scoring criteria should be handed out and discussed with the students. This could occur at the beginning of the course if a generic rubric is used for scoring all tasks or at the time the task is assigned. Discussing the scoring criteria with students helps them to understand the meaning of the scale, what is expected to receive the highest rating, and the importance of the task to their learning of the discipline.

In choosing dimensions for scoring an assessment task, focus on the intent in terms of instructional goals and the performance observable through the task. Two questions which help in choosing dimensions are 1) What are the criteria for selecting the samples that go into the portfolio and 2) What are the criteria for judging the quality of the samples? For example, if communication of statistical results is an important criteria, then it should be developed as a scoring dimension and its characteristics described. In a scoring rubric, there will be as many descriptions of the characteristics as there are points on the scoring scale. If you are rating the piece on a three point scale of excellent, acceptable, or poor, then in our example there would be three descriptions of the characteristics of communication (see Figure 1 at the end of this chapter for an example).

It is also helpful if you share examples of student work which exemplify the criteria for judging performance. The first time you use the assessment task this may not be possible but it is then that you can ask students if you might retain a copy of their work to serve as an exemplar. Over the years you may find it helpful to collect examples of both excellent and poor responses to share with your students. All exemplars should of course be used anonymously.

## PORTFOLIO SKETCH

### Introduction

In this section, the background information given thus far in the chapter is used to form a portfolio for a specific class. The portfolio sketch is an example of how a portfolio might be developed for a graduate level statistics curriculum. The sections match the important instructional targets or outcomes of the class. The computer programs and correct interpretation of results demonstrate mastery of the computer and statistical packages as well as an understanding of the results of running programs on real data. This section of the portfolio has additional value as a resource when students must analyze numerical data in future courses and on their thesis or dissertation.

The projects engage students in authentic, real-world applications of statistics. The corrected exams, on the other hand, measure the same knowledge in a more isolated, objective manner. The reflective journal is a measure of how well students have learned to analyze their strengths and weaknesses. Do they understand how they learn best and have they attempted to utilize

those methods more while applying skills to projects and exams? They are encouraged to write in the journal weekly. At a minimum, they are expected to write entries on their learning experience on the other sections of the portfolio. Some of the sections take just one or two weeks to complete, some take all semester, and some are group projects which are completed outside of class. The following sections represent the learning goals for the course.

| | |
|---|---|
| I | Mini Research Paper |
| II | Computer Programs and Interpretation of Output |
| III | Group Project |
| IV | Midterm Exams |
| V | Reflective Journal |

### *Before: Preparing students for the use of portfolios.*

At the beginning of the class, the portfolio is explained as a part of the course requirements. Many students are not familiar with portfolio assessment and even more have never been required to reflect on their learning and correct papers and exams after they have been graded. A lengthy discussion sets the expectation and purpose for these activities. I have found over years of using portfolio assessment that graduate students who are also fully employed professionals perform better with clearly delineated expectations. Therefore, the portfolio sections and criteria are established at the outset rather than allowing students to choose entries to the portfolio. A handout is distributed which outlines the contents and establishes the scoring rubric. The rubric gives the students a narrative description of the scoring criteria which communicates what they will have to do to receive the maximum, medium, and minimum number of points on the portfolio. Exemplars are distributed to reinforce the criteria and use of the rubric (see Lajoie, Chapter 14).

### *During: Issues in implementing a portfolio system.*

Often during the semester, students are given an opportunity to discuss and ask questions regarding the portfolio. Students who have not been journaling hear from their peers how this activity has helped them. Examples of a more productive use of study time and learning resulting from correcting projects and exams are motivating to students who have not been as successful in these activities. Students are asked to share with others in groups what they have in their portfolios to date.

Students are also encouraged to turn in their portfolios during the semester for feedback from the instructor. This informs both the teaching and learning process and helps students to understand that they are responsible for constructing their own meaning out of the course material. Writing encouraging comments as well as criticisms is important in order to motivate and keep the students actively involved in the learning process. The following sections have evolved over time and seem to work well in this graduate level statistics course.

*Section I: Mini Research Paper*. One of the first learning tasks I assign in my graduate statistics course is in the area of nonparametric statistics. As already mentioned, learning and assessment tasks should be parallel in content and context and in some cases a learning task may be extended into an assessment task. This learning task evolves into an assessment task and, as

part of the portfolio, becomes a baseline of comparison for later work as well as providing the first opportunity for reflection in the journal on the concept of statistical significance.

The task involves the student in choosing a problem or question which can be answered by collecting binomial data. The data must represent either yes/no, present/absent, success/failure or any other response which has only two possibilities such as heads and tails on a coin. Students formulate the question, collect a quick sample of data which should be obtainable in one to two hours, and then write an abbreviated research paper. I hand out several examples of the paper which includes an introduction of the problem, a null hypothesis, the methods used to collect the data, the findings section which consists of a visual representation of the data, and the conclusion.

The resulting paper as described above is presented in class and then handed in for initial feedback. The oral presentations can be done as a whole group in a small class or in groups of five to six in a larger class. I ask students to write a conclusion to their research although they are not trained to test for significance of the findings and the class discussion is guided into questioning the validity of their conclusions. Then I introduce a simple nonparametric statistical test, the binomial or sign test, and a six-step hypothesis testing framework. We work several sample problems together in class using student data. They learn how to make the determination of a significant difference between the two data points. The assessment task which goes into the portfolio is the research paper which they have edited and revised. This includes conducting the binomial test on their data and presenting the results in the findings section, writing a new conclusion, and editing the initial paper reflecting the comments I have made.

I begin the class with the mini research project to dispel the fear many doctoral students have that the intellectual demands of doing research are beyond them and most of all that they will not be able to understand and perform the statistical procedures necessary to complete the coursework. Through the mini research paper they discover the research process, the parts which logically present research to others, and the concept of significance.

*Section II: Computer Programs and Interpretation of Output.* During the semester students learn how to operate in the mainframe environment and run tests on two statistical analysis software packages, Statistical Analysis System (SAS) and Statistical Package for the Social Sciences (SPSSX). The second section in the portfolio contains SAS and SPSSX programs and listings. They learn to use SAS to run both descriptive and inferential statistics. Some of the programs they learn are procedures for means, correlation, analysis of variance, general linear models, and regression analysis. The number of programs taught on SPSSX is fewer. I find this package particularly good in describing data and presenting it in a way which can be transferred into a written document. Some of the procedures they learn on SPSSX are frequencies, tables, and reports.

The learning process is guided by several steps which individual students begin to omit as they become comfortable with the procedures. First they are introduced to the program in class both off and on the computer system. I hand out the printed program and explain each line in the program and what it tells the computer. We then calculate the most simple statistics by hand so that they know what the results of the program will look like. We then move to the computers and they learn to get on line on the mainframe and copy the prepared program off of my instructor disk onto their student disk and run the program. They then learn how to access the listing, the results of running the program, and print out both the program and the listing.

Each program is then practiced using different data sets which the students must enter into the program and run successfully. One of the programs which they have written and the resulting listing is put into the portfolio as an example of each procedure. When the course is complete they will have produced six to eight SAS and three SPSSX programs and listings.

Next I give them a research problem and some data and ask them to work in groups to decide what the hypothesis is and what analysis programs they need to run. They then work on the computers in groups to write the new programs, run them, interpret the listings and present the results in writing. This is practice for the group project.

*Section III: Group Project.* In the group project the students choose which type of design and statistical test they want to use. Each group writes a hypothetical problem, either collects or constructs the data, and chooses the type of analysis which is appropriate. They must make several decisions regarding the data set, what data to "collect," how to enter and label the data, and how to analyze the data. The entire study is then written up by the group and the results presented to the class. This group project goes into the portfolio.

*Section IV: Midterm Exams.* The midterm exams, of which there are two, are also placed in the portfolio. After the exam, the students are to analyze their own performance, correct wrong answers, and choose one area in which they need additional study. After the students have had a week or so to correct the exam, I review each response and answer any questions which remain. They enter into the portfolio some evidence of their study and the corrected responses to the questions on the exam which presented difficulty. In some cases, students may need to study more than one area of the exam in order to make the corrections. If they provide evidence of extensive study they earn bonus points on the portfolio which in turn may improve their grade.

*Section V: Reflective Journal.* One of the most important aspects of the portfolio is the reflective journal. As each assessment task is entered into the appropriate section of the portfolio, I ask them to analyze their work. Some of the questions I ask them to answer are 1) What did you have trouble with in producing this product? 2) How did you change the product after the initial feedback? 3) How could the product still be improved? and 4) What was the most important thing you learned in producing this product?

Additional questions for self-reflection on the group project include: Did you work with someone? How equitably did you contribute to the joint project? Did you help someone else? Who did the work on the computer? What problems did you encounter? How did you solve them? Did someone give you help?

The self-reflection required by this section involves recognizing the processes of metacognition which are stimulated. Bringing mental cognition out into the open through written reflection helps to teach students the mental processes they use effectively. Some of the thinking processes they should exhibit and recognize include analysis, synthesis, and evaluation and the application of these to the problem. Students' reflections can reveal much about the way they learn. By reviewing students' comments, the instructor may gain insights into the most effective and least effective teaching strategies they use. Here are a few examples which typify entries in student journals.

> At this point, had a glimmer the study was not addressing means so went back to my handouts and re-read chi square. Next tried to use a goodness of fit approach and realized attributes were not

dichotomous so my search continued. As I was flipping through notes, AHA, remember correlation diagram and decided this was a correlation problem.

Explaining out loud **why** ... listening to myself explain a topic or premise is a beneficial learning tool - no matter how time consuming. Teaching to the invisible student helps me clarify my knowledge or gaps of knowledge.

Correcting the tests is a good learning process. A good review of what missed and makes you go back and review those areas.

Use of a summary sheet acts like a self-quiz. By writing down what I think to be accurate in a visual table or format I can conceptualize better, check for accuracy easier, and reinforce progress to date!

Self reflection and correction of exam is an excellent learning tool. Rather than dismiss my exam as over I developed a deeper understanding of material. I will continue this approach after this class and have recommended it to others.

At this point, had a glimmer the study was not addressing means so went back to my handouts
Six-step hypothesis testing is a sequential process which keeps me on track and helps to define significance and draw conclusions... I have rewritten examples from class to practice this method. Another advantage to the 6-step approach is that it reduces my math anxiety.

Student comments emphasize the value of the reflective journaling. They tell their fellow students the value of this technique and continue to communicate its worth in office visits after the class is over. My comments on the journal focus on extending the students' thinking. This section of the portfolio is the most difficult to grade but probably of the most value to the student.

*After: Interpreting and using portfolio information.*

The students are encouraged to hand in the portfolio as soon as it is complete so that the grade and feedback may reach them before the final exam. They are also encouraged to make an appointment and discuss their learning process and progress on the material following the portfolio grading. The portfolio is awarded 300 total points which is about one third of the total for the course. In some cases, such as the first project and two midterm exams, the work received a grade when it was handed in originally. The revised version or corrected exam and the accompanying reflective journal entry is scored as part of the portfolio.

The points awarded in each section correspond to the scoring rubric. As described earlier, the rubric includes a description of the dimensions for scoring, the scale or points possible on each and the characteristics of the dimension expected at each level. Figure 1 (at the end of this chapter) provides an example of the general rubric used in scoring much of the portfolio. The dimensions in the rubric are statistical knowledge, strategic knowledge, and communication. The scoring rubric developed by QUASAR (Quantitative Understanding: Amplifying Student Achievement and Reasoning), a national project funded by the Ford Foundation, served as a guide for the development of the general rubric (see Lane, 1993).

## Summary

The overall purpose of the sketched portfolio is for the student to demonstrate competence as a researcher. The completed product illustrates what the student understands and is able to do in the area of statistics. It also indicates where more study or reflection is necessary in order to demonstrate competence. The reflective journal is a key component of the portfolio due to the conceptual nature of statistics. Often a clear understanding of the concept only requires more processing time on the part of the student and the journal requires active mental processing.

The elements described in the portfolio sketch are examples of what might be included. The elements should match the desired course outcomes or critical learnings. Some other options include asking students to analyze the methodology of a study published as a journal article, critique claims in a newspaper article using survey or other statistically based information, or create their own article for a student newsletter or write a conference paper. The portfolio can be extended and adapted to any educational setting or student population.

## IMPLICATIONS, CONCERNS AND BENEFITS

The complex practices and processes involved in portfolio assessment require a great deal of time and effort. Teachers must commit not only to creating learning activities, assessment tasks, and opportunities for reflection but must also be committed to using the results, being affected by them and by the insights into their teaching which result. There is also time involved in grading the portfolio. If you don't have time to read the student's work then don't choose portfolio assessment. If you have an assistant or grader, it is best to have the person grade the exam and grade the portfolios yourself. Bias is a difficulty in grading this type of assessment. If the class is small, you will know each person's work and their progress throughout the course. This means, as an instructor, you must be aware of any bias you have concerning individuals and take precautions that your bias does not enter into the grading process. One means of guarding against bias is by grading each portfolio and then going back through each one a second time to check for consistency in the awarding of points.

When students' knowledge results from an active transaction with the information being learned, it is best if portfolio entries are made throughout the year. In this way, assessments become embedded in the topics of study and reinforce hands-on teaching and learning (Shavelson & Baxter, 1992). This gives teachers and students opportunities for reflection and knowledge of what needs to be reinforced or reviewed throughout the course. The multi-source nature of portfolio assessment affords students more opportunities to demonstrate their knowledge and to show improvement. As teachers, we must model the complex habits of mind and ways of processing information which we ask of our students and ask for feedback from students on new activities which evolve from the process. This self-analysis has the potential to improve the strategies used in the classroom as well as the tasks. It gives the teacher more formative information on the effect of instruction on learning. As tasks are completed, the teacher can analyze strengths and weaknesses in teaching and make changes which will increase learning.

Another benefit of portfolio development is the dialogue which takes place between the students and the teacher. This dialogue begins with a discussion of what will be included in the portfolio as evidence of concept mastery. It may be an oral dialogue or a written one such as in student journal writing and teacher response. When students begin to write reflectively, there

always seems to be discussion in class of what and how to write. The teacher needs to reinforce the part this process plays in students taking responsibility for their own learning. Students who by this practice continue to expand their understanding and construction of a cognitive web of statistical knowledge benefit greatly. The addition of the reflective process in the form of journaling facilitates this construction of knowledge. Knowing when to step in and redirect a student' s efforts and when to step back and allow for the process of discovery takes experience with this type of discourse (Zessoules & Gardner, 1991).

Portfolio use can extend learning into the assessment process. Portfolio assessment is also extremely flexible and may be adapted for a wide variety of settings including K-12, post-secondary education, and graduate programs. The issues and concerns are different depending on the setting and use. In K-12 systems, there are many schools and school districts who have adopted portfolio assessment in various subject areas (see Brandt, 1992). Most experiences have been positive and participants have found that assessment projects engage both teachers and students in reflection on the learning process. In some cases, portfolios have motivated students and given their writing purpose. One teacher concluded that long-term situational math problems included in the portfolio were of the most lasting importance to students (Knight, 1992). The Portfolio Assessment Clearinghouse publishes the Portfolio News quarterly which includes descriptions of portfolio projects, articles on how and why to use portfolios, and reviews of the literature (Cooper & Davies, 1990 to present).

Portfolios have been developed at the state level. A 1990 survey by the Center for Research on Evaluation, Standards, and Student Testing at UCLA found that nearly half of state testing programs either had alternative assessments in place, were planning to implement them, or were actively exploring the concept. Alaska has encouraged innovative assessment projects and has assisted a number of districts in the area of portfolios. Several states are using portfolios for documenting work in mathematics and writing. Vermont and Kentucky have had 4th and 8th grade writing and mathematics portfolios since 1992. California initiated the use of portfolios as part of their performance assessment project in the 1993-1994 school year. These large-scale state efforts have encountered some resistance both from factions of the community against performance assessment and from teachers.

Some lessons have emerged from the large-scale portfolio assessment programs in Vermont and California. In a report of the pilot year of Vermont's Mathematics Portfolio Assessment Program (Vermont Department of Education, 1991), several suggestions for improvement of the system were made. Among those listed is the need for additional professional development and materials to be committed to teacher training. A recent Rand study found that, as a result of the training provided since 1991, teachers are teaching more problem solving but that there is still a need for more training (S. Rigney, personal communication, March 8, 1995). Due to some additional training on the structure and contents of the portfolio, the number of scorable portfolios has gone up considerably from the 83% of 4th grade and 58% of 8th grade portfolios reported in the earlier state publication. The inter-rater reliability of scoring remains an issue although it has improved over the four years of the program. It is now reported to measure r=.80 in mathematics and an r=.70 in writing. The Vermont program is continuing and the teachers and other participants seem to be committed. There is an understanding that it involves systemic change and as such takes time to implement successfully (S. Rigney, personal communication, March 8, 1995). The California project had only one sampling and scoring session prior to the discontinuation of the program in December, 1994. The issues which emerged from studying the results were 1) a question of whose work was represented by the portfolio, 2) the consistency of

the pieces included, and 3) the reliability of the scoring (B. Thomas, personal communication, March 14, 1995).

The "whose work is it" question is of concern with all assessments completed over time. This question is not unique to assessment; any assignment done outside the classroom is open to the same criticism. There is also a controversy surrounding group tasks. If we believe in the merit and benefits of group efforts in learning, then this issue of "who's work" should not eliminate the inclusion of group assessment tasks in the portfolio. If the critical elements of cooperative learning are included in the approach to using groups, then the individual would still be held accountable for learning (Johnson, Johnson, & Smith, 1991). The important point is for raters to know the conditions under which the work was done. If you as the teacher act as the rater, then you would have this context information and the ratings would more likely be reliable.

Interest in alternative assessment is growing although the form the new assessment will take and many issues in the development are still being debated. Each instrument needs to be thoughtfully crafted and requires a specially developed or adapted scoring method. Portfolio assessment is just one form of assessment based on demonstrating outcomes. Because it is multi-source, some evidence of reliability is evident. The inclusion of both traditional and authentic sources of data complete the picture of the learner and help to direct learning to meet student needs. The portfolio guides the student through powerful, on-going assessment and provides a record of progress and a resource for future learning.

**General Rubric**

| **Dimensions** | **Score Level and Points Awarded** | **Characteristics of Response** |
|---|---|---|
| | *Level 3*<br>Excellent<br>Total points awarded project may be 50 to 60 | Shows understanding of the problem's statistical concepts; uses appropriate terminology and notations; executes computations completely and correctly. |
| **Statistical Knowledge** | *Level 2*<br>Acceptable<br>Total points awarded project may be 39 to 49 | Shows understanding of some of the problem's statistical concepts; uses nearly correct terminology and notations; may contain computational errors. |
| | *Level 1*<br>Poor<br>Total points awarded project may be 0 to 38 | Shows very limited or no understanding of the problem's statistical concepts; may misuse or fail to use terminology; makes major computational errors. |
| | *Level 3*<br>Excellent<br>Total points awarded project may be 50 to 60 | Identifies all the important elements of the problem and shows understanding of the relationships between them; gives an appropriate and systematic strategy for solving the problem; gives clear evidence of a solution process which is correct. |

| | | |
|---|---|---|
| **Strategic Knowledge** | *Level 2*<br>Acceptable<br>Total points awarded project may be 39 to 49 | Identifies some of the important elements of the problem and shows some understanding of the relationships between them; gives evidence of a solution process which may be incomplete. |
| | *Level 1*<br>Poor<br>Total points awarded project may be 0 to 38 | Fails to identify important elements or places too much emphasis on unimportant elements; may reflect an inappropriate strategy for solving the problem; solution process may be missing, difficult to identify or completely unsystematic. |
| | *Level 3*<br>Excellent<br>Total points awarded project may be 0 to 38 | Gives a complete response with a clear, concise explanation or description; presents supporting arguments which are logical; may include an appropriate diagram and examples. |
| **Communication** | *Level 2*<br>Acceptable<br>Total points awarded project may be 39 to 49 | Gives an explanation or description which may be somewhat ambiguous or unclear; may include a diagram which is flawed; arguments may be incomplete or may be based on an unsound premise. |
| | *Level 1*<br>Poor<br>Total points awarded may be 0 to 38 | May have some satisfactory elements but may fail to complete or may omit significant parts of the problem in the explanation or description; may include a diagram which incorrectly represents the problem, is difficult to interpret, or no diagram. |

**Figure 1: Portfolio Scoring Rubric**