Journal of Personnel Evaluation in Education 15:2, 121–133, 2001

Presenting Complex Teacher Evaluation Data: Advantages of Dossier Organization Techniques Over Portfolios

KENNETH D. PETERSON*
Portland State University

DANNELLE STEVENS
Portland State University

CAROL MACK
Portland State University

Abstract

The development of extensive and authentic teacher evaluation raises questions about how best to organize and present the increased amount and variety of assessment materials. Portfolios have been suggested for teacher evaluation because they are well suited to capture the complexities of teaching. However, portfolios 1) are difficult to judge, 2) are difficult to archive, 3) lack key information, 4) miss excellent teachers with simple materials, 5) rarely plan for audiences, 6) distort when required, and 7) bring conflicts of interest. Teacher dossiers are compressed collections of objective data which are easier to judge, demonstrably reliable, and cost-effective.

Keywords: teacher portfolio, performance evaluation, documentation

Most current school teacher evaluation data come from an administrator classroom visit and a report form (Loup et al., 1996). This strategy often uses rather narrow indicators of teacher quality such as a checklist rating of a segment (e.g., 30 minutes) of classroom observation, casual review of available lesson plans, and incorporation of haphazard hearsay general impressions of fellow teachers, parents, and students. However, teacher evaluation can be made more useful and effective by collecting a greater amount and variety of information about what teachers do and how they affect the people with whom they work (Peterson, 2000; Wolf et al., 1997). This new emphasis in teacher evaluation is to base judgments on a variety of data sources of the genuine teaching work and performance of the teacher in the form of test scores, client surveys, peer reviews, evidence of activity, materials, work samples, and products. This more comprehensive and accurate assessment data for teachers can result in ''authentic'' (Perrone, 1991) evaluation by being more (a) realistic in content and performance; (b) comprehensive in scope and

*Correspondence should be addressed to: Kenneth D. Peterson, School of Education, Portland State University, Portland, Oregon 97207. E-mail petersk@mail.pdx.edu

inclusion about what teachers actually do; and (c) accurate by replacing second-hand administrator impressions with primary data such as client surveys and test scores.

The development of extensive and authentic teacher evaluation raises questions about how best to organize and present the increased amount and variety of assessment materials. Two such methods that have been described in the literature are the portfolio and the professional dossier. Portfolios have been suggested for teacher evaluation because they are well suited to capture the complexities of teaching (Sergiovanni, 1977; Wolf, 1991, 1996; Wolf et al., 1995). Bird (1990) described portfolios as a way to see school teaching as ''a form of expression, a humane project, an evolving state of affairs, and a situated accomplishment over time'' (p. 249). Just as other kinds of portfolios are highly specialized (e.g., program portfolios, Ponzio et al., 1995), the unique possibilities of teacher portfolios need careful consideration before widespread adoption.

In the 1980s the Utah Teacher Evaluation Project (UTEP) (Peterson & Kauchak, 1982) faced the challenge of changing educator evaluation based on administrator reports, which rarely discriminate among teachers, to a new approach that used the widest range of data available for career ladder promotions. We took the best ideas available from portfolio strategies, but had the realistic problem of how to use and judge the data that teachers would choose to assemble. We found a competitive modification of many portfolio ideas in the form of the teacher dossier (Scriven, 1973a, 1973b), an alternative to professional portfolios for teacher evaluation. Teacher dossiers are a more compressed and restricted organization technique. This study will report the key features of portfolio documentation strategies, highlight distinct disadvantages of portfolios found in the UTEP field implementation, describe the distinctions of dossier techniques, and recommended dossiers over portfolios for school teacher evaluation because of their greater utility for review.

Selected Literature on Teacher Portfolios

A number of writers have contributed to the literature on designing and using teacher portfolios. These ideas and experience are important to this study because they highlight the potentials and problems of portfolio documentation of teacher performance which lead to our development of professional dossiers for teacher evaluation.

Brauchle et al. (1989) reported two years of field data on the use of portfolios to document student performance data. In this system, teachers in a career ladder program were given the responsibility for selecting and presenting their evidence of improvement of pupil achievement and attitude toward learning. Difficulties with this process were numerous. Teachers were not able to document student achievement: ''. . . many . . . teachers appeared not even to know how to begin'' (p. 28). Second, fair treatment across subject areas and grade levels was a significant problem. Third, ''a flock of logistical problems remains to be solved'' (p. 28). These include the finding that teachers collect and present great masses of data of questionable quality and relevance, with great expenditures of time. Questions of authenticity and responsibility for

results were frequent. The authors concluded that ''it would probably be easier and less expensive to give the students a standardized test and compare student average scores. [However,] . . . the ability to provide a flexible format for evaluation targeted to the specific instructional goals of each teacher . . . is worth the added effort'' (p. 28).

Bird (1990) discussed four central purposes of teacher portfolios: selection and hiring, professional development, school reform, and professionalization. He listed possible contents of portfolios: photocopies, photographs, observation reports, video tapes, and other unique evidence. Teachers should consider the meaning and significance of possible entries that they hold in common with the persons who use the portfolio: ''The initial task is to organize . . . a penetrating and useful conversation about schoolteaching'' (pp. 248–9). Bird analyzed content organization in a matrix of five task areas (teaching a class, planning and preparation, student evaluation, professional exchange, community exchange) with four dimensions of responsibility, subject matter, individual students, and class organization. Clearly, this 20-cell matrix represents much complexity to document. Teachers and their audiences must be careful and explicit in why, how, and when teachers produce portfolios.

Seldin (1997) reported that portfolios are well developed in postsecondary teacher evaluation where they are found in over 1,000 institutions. He described portfolios as '' . . . *selected information* on teaching activities and *solid evidence* of their effectiveness'' (p. 2). The purposes of a portfolio in higher education are to present hard evidence and specific data for promotion and tenure decisions, gather materials for self reflection, see how teaching evolves over time, share experience with newer colleagues, seek awards or grants, and leave a written legacy for future teachers in the department. Seldin recommended sections in the portfolio consisting of materials from oneself (responsibilities, context), materials from others (colleague statements), products of teaching/learning (test scores, work samples), and unique contents (videotapes, professional activity). Seldin acknowledged that busy reviewers may skim parts of portfolios rather than read them in depth, and that ''unexplained evidence is difficult for readers to understand and interpret'' (p. 23).

Wolf et al. (1996) reported a study of portfolios for summative decisions in a pay-for-performance system in which volunteer teachers could elect evaluation for a one year, $1,000 stipend and recognition as an Outstanding Teacher. Formative professional development was a secondary benefit of the program. The portfolio format was a notebook with seven components: resume, philosophical statement, three brief commentaries on three scoring categories, six artifacts with rationale for selection, reflections on peer survey (four chosen by teacher) and client survey (20 students chosen each by teacher and administrator), recent evaluation report, and a self-evaluation. The portfolio was judged by an administrator, from the candidate's building, who received one day of training. Three areas of review were Assessment and Instruction, Content and Pedagogy, and Collaboration and Partnership. Three findings were possible: Outstanding, Excellent with no reward, and Not Meeting Criteria. Judgment could include data or perspectives outside of the portfolio, and was aided with suggested steps for review. Of the 829 eligible teachers in the District, 266 submitted portfolios and 236 (88.7 per cent) were granted the

reward. The middle category of Excellent was not used by the administrators. The authors concluded that, in general, the program was well received (it remained in the negotiated contract). Specific benefits noted included clarification of goals for teachers and for administrators increased information about the good work going on in the building. Suggested changes for the future were to add a second administrator view, more emphasis on pupil achievement data and school improvement, and to minimize the negative effects on the teachers found not to meet criteria (those cases could be ''devastating''). The authors were ''enthusiastic'' about the program and reported ''important contributions to teaching and learning'' (p. 286).

Wolf et al. (1997) summarized state of the art advice for teacher portfolios. They defined a teacher portfolio as '' . . . the structured documentary history of a carefully selected set of coached or mentored accomplishments, substantiated by samples of student work, and fully realized only through reflective writing, deliberation, and serious conversation'' (p. 195). This documentation may include resumes, letters of recommendation, client surveys, lesson plans, and evidence of student learning tied to established standards. They warned against mere collections or scrapbooks of classroom life with ''affectionate notes from students and parents.'' While they recognize that evaluation of portfolios is a ''daunting task,'' they recommend as helpful remedies (a) identification of content and performance standards; (b) specifications for portfolio construction; and (c) design of an efficient evaluation system.

Study Design: The Utah Teacher Evaluation Project

In the 1980s the Utah Teacher Evaluation Project designed and studied teacher evaluation systems in five school districts as a part of the Utah Teacher Career Ladders program set by the State Legislature (Peterson & Kauchak, 1982, 1983; Peterson et al., 1986; Peterson & Mitchell, 1985). The UTEP was centered at the University of Utah, but engaged the efforts of planners in the school districts and representatives from many other agencies in Utah including the Utah Education Association, state Parent Teachers Association, League of Women Voters, Legislative Research Analyst's Office, Governor's Educational Advisor, planners from several large businesses, and Guadaloupe Center (private, non-profit ethnic support organization).

The central effort of the UTEP was to establish teacher evaluation procedures for promotion to a position on a career ladder scheme based upon demonstrated quality performance, preparation, and outcomes. Multiple data sources were established, including student and parent surveys, pupil achievement data, teacher test scores, peer review of teacher materials, improved administrator reports, systematic observation (by observers outside of the local district), documentation of professional activity, and development of unique evidence by individual teachers (McCarthey & Peterson, 1987; Peterson & Kauchak, 1983). Each data source was selected by the teacher, but collected and reported by an external UTEP Evaluation Unit distinct from the district. These data were organized by individual teachers in varying combinations to make the best case for that person's promotion. District-wide, teacher-dominated panels were established to

make the summative judgments about promotion and granting of salary increases (ranging from $750 to $2,000 per year for five to seven years). The eight-person panels of teachers, administrators, and parents were selected to not personally know the candidates.

Small scale pilot studies of this system called for reports of teachers on their data sources. The early offerings were teacher portfolios with widely varying collections of materials and self-reports. Problems with interpreting these collections led to systematic identification of seven consistent problems, and subsequent development of teacher dossiers as a desirable replacement of the initial portfolio system. Early in the UTEP experience guidelines (Peterson, 1988) for professional dossiers replaced options for portfolio presentation. The following sections of this paper detail the seven problems, and the resulting dossier system derived from the UTEP.

Findings: Disadvantages of Portfolios for Teacher Evaluation

Seven serious problems with teacher portfolios which preclude their use for summative evaluation are discussed in this section. These problems mean that portfolios should be used as a single *optional* data source for teachers, judged by peers who produce a one to two page summary, and included with other data sources in a summary professional dossier.

Scoring Difficulties

While they are rich data sources, portfolios are notoriously difficult to use for judgments (Brauchle et al., 1989; Peterson, 1995; Wheeler, 1994; Wolf, 1991; Wolf et al., 1997). The open-ended nature of portfolios is an advantage for capturing the variety of teaching excellence; but it is a disadvantage for deciding about quality. Portfolio structure allows individuals to custom present materials, and it allows for individuality and documentation of specific examples. However, non-uniformity means difficulty in making comparisons or for judging overall adequacy. Unorganized collections of teacher and student artifacts provide few inherent standards for comparison or judgment.

Physical Difficulties

Teacher portfolios are bulky. Unless limited by ground rules, portfolios may grow to include artists' containers of poster sized materials, boxes of papers, large collections of student work samples, and classroom realia (we received a student-made corn tortilla in one). Portfolios are difficult to store, and take up much room. Archived portfolios are little used, except for people who search them for ideas on how to put their own portfolios together, rarely to improve teaching. A new set of problems begins if portfolios are not archived, but left with teachers. First, they are inaccessible for reference by others. Second, they are subject to raids by the teacher for materials and examples to use in the

classroom, or to loan to others who are assembling their portfolios. It is rare for portfolios to stay intact if they are kept by the teachers who produce them.

Lack of Important Information

Portfolios most often leave out a great number of needed perspectives for judgment of teacher quality. For example, portfolios may show sample results, but not overall levels of pupil achievement. Portfolios rarely include systematic summaries of student and parent opinions about the impact of the materials on display. The reactions of peer colleagues are important guides to the value of portfolio contents, but rarely are included. Also, the preparation and out-of-class duties of a teacher often are significant concerns for summative evaluation. Of course, these perspectives can be included in a portfolio, but their presence often only complicates the organization.

Some Teachers are Good for Reasons Other than Materials

Teacher portfolios may seriously underplay the strengths of teachers whose quality is *not* evident in materials or student products. Good teachers may be good because of personal interactions, persistence, hard work, inspiration, pupil achievement in standardized dimensions, productive routines using few materials, a richly imaginative classroom, success in encouraging a divergence of student production, or educational payoffs with only long term visibility. All of these teacher dimensions may be under emphasized in collections of teacher materials.

Audiences Rarely are Planned for

The first part of portfolio design is a set of good procedures to accumulate a physical record. The second part of the design (quite often neglected in educational practice) is to plan for the review, inspection, or dissemination audiences. Creating a good portfolio is only half the task; the second half is to plan for the person or groups to view, use, and comment on the work. Portfolios are time consuming and difficult to produce. If producers do not have appreciative audiences, many hard working teachers become disappointed and resentful.

Required Use Distorts Practice

The mandated use of teacher portfolios in summative evaluation distorts both the evidence and process of judgment. Summative uses reward portfolio producers, not necessarily the best teachers. Portfolios are a secondary or even a tertiary kind of evidence (relative to classroom performance or pupil achievement). An evaluation system that places a

premium on portfolios soon creates an industry of portfolio assembly far beyond the intended authentic samples of teacher work.

Self-evaluation and Reflection are a Serious Conflict of Interest with Summative Teacher Evaluation

Self-evaluation and reflection are important and necessary professional activities. Every good teacher is expected to practice these activities in some form and emphasis. Good formative evaluation necessarily takes the individual teacher's perspectives and judgments into account. However, in a summative evaluation procedure, the individual teacher is placed in a conflict of interest situation where her interests (status, prestige, advancement, security) must be juxtaposed with those of her clients (students, parents, administrators). Reflective statements invite judgments based on rhetorical skills quite independent of classroom quality. The purposes and demands of summative teacher evaluation call for a separation of self-report and distinct evidence of preparation, performance, and outcome. The heavy emphasis on self-evaluation of most portfolio systems is a problem for summative teacher evaluation.

Teacher Dossiers

A *dossier* is a collection of documents related to a specific matter. A teacher professional dossier has the purpose of attesting to teacher quality (Scriven, 1973a, 1973b). Thorough and useful teacher dossiers can be limited to 12–15 pages in length while including many of the benefits of portfolio documentation (Peterson, 1984, 1988, 1990, 2000). The distinction between portfolios and dossiers is non-trivial. Dossiers are much more compact, processed, and usable for judges of teacher quality. Central to the distinction between portfolios and dossiers is Scriven's (1973a) idea of *compression* in educational evaluation. The voluminous data that can be gathered in authentic evaluations require treatment before judges can do their work. Compression is a process in which evaluation data are summarized, processed for key information, and subjected to prior sub-judgments of specialized reviewers. One necessary preparation for teacher data is a *credentialling* of standards and claims in which expert views of the quality and importance of goals, methods, and results are documented. Compression greatly reduces size; for example, a one page peer review of materials summary can represent a box of materials that contains over 500 pages of plans, procedures and products! (McCarthey & Peterson, 1987). Lists of contents for three sample teacher dossiers are presented in table 1.

A central feature of evaluation is to examine collections of data and other information to make value judgments. Scriven (1967, 1973a, 1973b) described evaluation as a process to ascribe merit and worth to some product, preparation, performance, procedure, or outcome. There is quality in and of itself (merit), and the value in relation to some audience (worth). The judgment part of evaluation is to determine this merit and worth in relation to a specific audience, as near to their terms and values as possible. Teacher

Table 1. Contents of Three Teachers' Professional Dossiers (Peterson, 1995).

TEACHER A
Documentation of Professional Activity—2 pages
Peer Review summary reports (1984, 1991)—2 pages
Student Report data* (1984–1994)—2 pages
Administrator Reports (copied, reduced 2/page)—4 pages
Parent Survey data* (1984–1994)—2 pages
TEACHER B
Student Achievement data (alternate years after 1985)—3 pages
Teacher Test Scores Report—1 page
Documentation of Professional Activity—2 pages
Systematic Observation Report—3 pages
Administrator Reports—1 page summary
Extended Parent Comments (reduced)—4 pages, with peer comments; Evaluation Unit description of comment selection
TEACHER C
Administrator Reports—4 pages
Annual Reports of Community Art Festival (reduced; 1980–present)—6 pages; Created by teacher C; Student community learning projects
Student Report data (1984–present)—3 pages
Parent Reports data (1992 focus group; alternate years after 1988)—2 pages

Note. *1984–1987, 1989, 1991, 1993.

dossiers present the collection of data and other information to the groups or individuals who make the judgments about merit and worth.

Teacher dossiers can be reliably used for summative judgments about teacher merit and worth. Peterson (1988) reported defensible reliability with dossiers used to make decisions about teacher promotion in the UTEP career ladder system. In this study, 12 teacher dossiers were judged by 26 career ladder panelists, including teachers, administrators, university faculty, and nursing educators. The dossiers in this study had a mean length of 12.7 pages and averaged 4.75 variable data sources. Panelists showed a 90.4 per cent agreement level on the 312 promotion judgments. The mean rank order correlation coefficient among judges (Spearman's rho) was 0.729. Cohen's Kappa for agreement on promotion was 0.808.

In another study of UTEP procedures (Peterson, 1990), dossiers were reliably used to make simulated rankings of teacher performance by a computer expert program. The simulation was only for the purposes of better understanding human judgments about teacher data—no ranking purpose was suggested for actual teacher settings. Dossiers presented sufficient data in an organized fashion to be used to respond to a bank of 111 questions on teacher quality. The computer expert system was able to categorize (promote or not promote) and rank the dossiers as well as the top person in the middle third of the group of human judges. Dossiers were described as valuable for increased understanding about how judges use complex teacher evaluation data.

Authentic assessment, as represented in panel reviews of teacher professional dossiers, can be used for decision making, public relations, teacher education program designers, and other external audiences. Thus, dossiers are a valuable asset for individual teachers, school districts, and others interested in teacher quality.

Guidelines for Dossier Assembly

Each district should establish ground rules for dossier assembly and use (Peterson, 1988, 1990, 1995). While teachers and dossiers should show individuality, personal strength, and authentic views, some guidelines and ground rules make them more usable and fair. Ground rules also give structure for teachers who need it. Ground rules may address size of dossier, information required for credibility, protection of due process, and expectations for performance. Guidelines should also be established for numbers of data sources, page size, and number of pages.

Credibility depends on confidence that dossiers contain honest information, and that changes and alterations are not made to falsely enhance them. Ground rules can be agreed upon to provide security to dossier storage. Storage should allow updates and protect confidentiality, but maintain accuracy of contents. It is necessary to have a safe copy developed by the district Evaluation Unit, and to have teachers keep a working copy.

Due process in handling professional dossiers is important. While safeguards are afforded by involving neutral participation of other people (district forms are used, independent clerks gather and score surveys, claims for professional activity are backed by letters, transcripts, testimonies in a teacher-held file), professional dossiers are voluntary and ultimately the property of the teacher. They are *not* to be considered personnel files or employment files, which are the property of the school district. The safeguards are to insure confidence, not to restrict the rights of teachers to construct and maintain an accurate record of their work.

Ground rules can assure technical adequacy. The overall ground rule is that each data source has guidelines that must be respected and documented in each case. For example, some return rates on parent surveys can be inadequate. Particularly at the high school level, some good teachers may have less than 40 per cent of mail surveys returned (Peterson, 1989b). This does not necessarily mean that the teacher practice is inadequate, only that the data gathering technique is. Ground rules that establish lower limits for data in dossiers insure technical accuracy. In the example of parent surveys, return rates must be clearly shown. District averages (obtained from the Evaluation Unit) indicate to the teacher and reviewers compliance with this guideline.

Finally, district ground rules can help to give perspectives on significance of reported values. For example, the absolute value of a global item on a student survey needs the perspective of district norms. Guidelines for value may be established, for example, ''good practice is indicated by averages at or above *1.5 standard deviations below the district mean*.'' This guideline gives perspective on values as teachers gather data and summative judges review dossiers.

Costs of Dossiers

Dollar and time costs of dossier preparation and maintenance are defensible; however, they need to be openly discussed and agreed upon (Peterson, 1989a). One dossier cost is personnel time; teachers need time to plan, talk with other teachers, research possibilities, and maintain their records. Evaluation Unit persons need time to file, review, and audit dossiers stored with the district. Another dossier cost is production, including duplication of papers for the dossier itself and for the teacher's back up file of transcripts, letters, test score reports, and testimonials. Organizing and presenting materials requires purchase of folders or dividers. Computer disks that contain teacher data may be an additional expense. While these costs may seem trivial, they should be openly presented and valued whether they are paid for by teachers or the district. Recognition and acknowledgment are key elements to the culture of the district, and the resulting respect for the evaluation system. The sociological value of district support far exceeds the monetary value (Johnson, 1990; Lortie, 1975; Peterson, 2000).

While dossiers are the initiative and property of teachers, school districts have interests in encouraging and having high quality teacher professional dossiers. In order to get good dossiers, districts have options about responsibilities for funding. Some options are: (1) district pays all assembly and maintenance costs; (2) district pays all dossier costs for the first dossier, teacher maintains; (3) district give supplies; (4) district gives supplies at a discount (bulk purchase); (5) teacher pays all assembly and maintenance costs; and (6) district schedules time for dossier work vs. teachers make time.

Summary Discussion: A Comparison of Portfolios and Dossiers

This discussion has described portfolios and dossiers as distinct techniques for presenting complex teacher evaluation data. Figure 1 depicts key features of this contrast. Our main purpose in making these distinctions is to think through issues concerning the improvement of personnel evaluation by including more complex data than are used in most current practices.

Teacher professional portfolios and dossiers share some characteristics for evaluation purposes. Each technique provides a kind of valuable process and information for teachers and those who evaluate them. Both emphasize choice by the teacher in content, organization, and emphasis. Each allows presentation of complicated information of preparation, practice, and outcomes. Customization is permitted by both documentation techniques. Important adjustments to the contexts and resources of performance are possible with either approach. Both portfolios and dossiers allow inclusion of a wide variety of teacher evaluation techniques such as judgment-based (Popham, 1988), system-wide standardized testing (Sanders & Horn, 1995), duties-based (Scriven, 1988), and performance assessment (Perrone, 1991).

However, since teacher portfolios currently enjoy strong advocacy, it is important to consider some of their limitations in light of a viable alternative. As many writers have noted, the *purpose* for which portfolios are used is very important. Most advocates

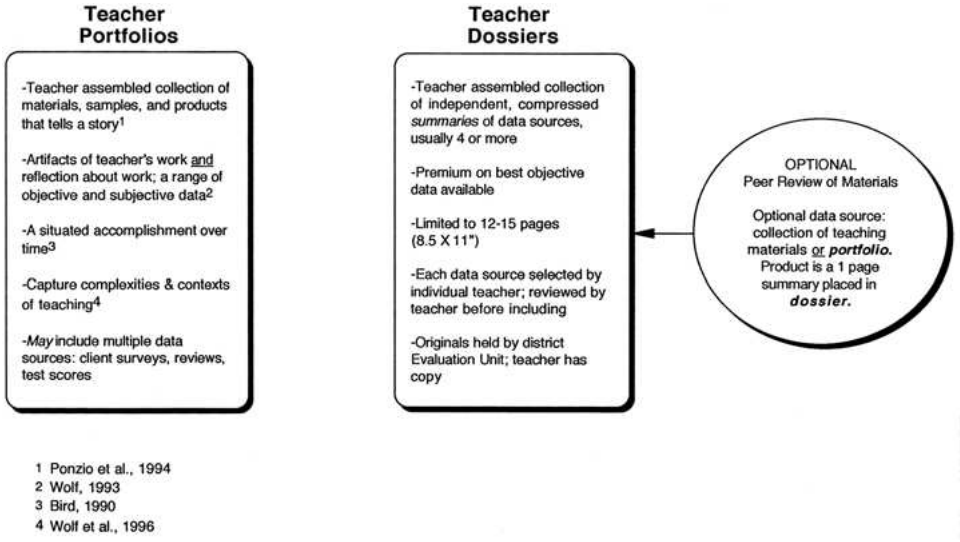

Figure 1. Comparison of teacher professional portfolios and dossiers (Peterson, 2000).

describe the strong formative values of portfolios. However, our experience with both portfolios and dossiers leads us to recommend certain considerations for summative teacher evaluation in which judgments will be made for retention, tenure, leadership appointment, or comparative performance recognition.

For these decision-making, summative functions of teacher evaluation dossiers present important improvements over portfolios. Compression of data, for example, presenting years of pupil surveys or extensive teacher test results on single pages, makes the work of a summative judge (or, panel) much more feasible. Pre-processing data, for example using a one page summary judgment of three peer teachers, rather than reviewing several boxes of teacher materials (student work samples, instructional materials, assignments, media, schedules) enables informed decision making. Teacher portfolios can be encorporated into dossiers, if (a) the individual teacher claims that the dossier review is valid and reliable; and (b) the value of the portfolio is described by peer teachers (not in a professional or social relationship with the candidate) on a page or two in the dossier. The emphasis of dossier techniques on the search for the best objective evidence available, rather than a global portrait of that practice helps with summative decisions. Reducing the emphasis on the perspectives of the teacher in terms of goals or values, makes comparisons with standards and summative decisions more possible. Down playing teacher rhetoric by eliminating goal and reflection statements relies on actual preparation, performance, and outcomes. Dossier focus on the purpose of presenting the comparative or *hard data* case for teacher quality most often better supports summative decisions. Finally, the smaller size of dossiers eases the logistics of storage and reproduction. These central reasons constitute a case for teacher evaluators to continue to think through the promises and problems of documenting complex teacher data organization techniques.

References

Bird, T. (1990). The Schoolteacher's Portfolio: An Essay on Possibilities. In J. Millman and L. Darling-Hammond (Eds.), *The New Handbook of Teacher Evaluation: Assessing Elementary and Secondary Teachers*. Newbury Park, CA: Sage, pp. 241–256.

Brauchle, P., Mclarty, J., & Parker, J. (1989). A Portfolio Approach to Using Student Performance Data to Measure Teacher Effectiveness. *Journal of Personnel Evaluation in Education*, *3*, 17–30.

Johnson, S.M. (1990). *Teachers at Work: Achieving Success in our Schools*. New York: Basic Books.

Lortie, D.C. (1975). *Schoolteacher: A Sociological Study*. Chicago, IL: University of Chicago Press.

Loup, K., Garland, J., Ellett, C., & Rugutt, J. (1996). Ten Years Later: Findings From a Replication of a Study of Teacher Evaluation Practices in our 100 Largest School Districts. *Journal of Personnel Evaluation in Education*, *10*, 203–226.

McCarthey, S.J., & Peterson, K.D. (1987). Peer Review of Materials in Public School Teacher Evaluation. *Journal of Personnel Evaluation in Education*, *1*, 285–293.

Perrone, V. (Ed.) (1991). *Expanding Student Assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

Peterson, K. (1984). Methodological Problems in Teacher Evaluation. *Journal of Research and Development in Education*, *17*(4), 62–70.

Peterson, K., & Kauchak, D. (1982). *Teacher Evaluation: Perspectives, Practices and Promises*. ERIC Document Reproduction Service ED 233 996. Salt Lake City, UT: University of Utah.

Peterson, K., Kauchak, D. Mitchell, A. McCarthey, S., & Stevens, D. (1986). *Utah Teacher Evaluation Project: The Park City Career Ladder Design*. ED 265 143. Salt Lake City, UT: University of Utah.

Peterson, K.D. (1988). Reliability of Panel Judgments for Promotion in a School Teacher Career Ladder System. *Journal of Research and Development in Education*, *21*(4), 95–99.

Peterson, K.D. (1989a). Costs of School Teacher Evaluation in a Career Ladder System. *Journal of Research and Development in Education*, *22*(2), 30–36.

Peterson, K.D. (1989b). Parent Surveys for School Teacher Evaluation. *Journal of Personnel Evaluation in Education*, *2*, 309–319.

Peterson, K.D. (1990). DOSSIER: A Computer Expert System Simulation of Professional Judgments on Schoolteacher Promotion. *Journal of Educational Research*, *83*, 134–139.

Peterson, K.D. (2000). *Teacher Evaluation: A Comprehensive Guide to New Directions and Practices* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Peterson, K., Kauchak, D. Mitchell, A. McCarthey, S., & Stevens, D. (1986). *Utah Teacher Evaluation Project: The Park City Career Ladder Design*. ED 265 143. Salt Lake City, UT: University of Utah.

Peterson, K., Wilson, J., Ulmer, D., Franklin, N., Mitchell, T., & Winget, L. (1984). *Teacher Career Ladders in Utah: Perspectives on Early Development*. ED 246 051. Salt Lake City, UT: University of Utah.

Peterson, K., & Kauchak, D. (1982). *Teacher Evaluation: Perspectives, Practices and Promises*. ERIC Document Reproduction Service ED 233 996. Salt Lake City, UT: University of Utah.

Peterson, K., & Kauchak, D. (1983). *Progress in Development of Lines of Evidence for Teacher Evaluation*. ERIC Document Reproduction Service ED 228 228. Salt Lake City, UT: University of Utah.

Peterson, K., & Mitchell, A. (1985). Teacher Controlled Evaluation in a Career Ladder Program. *Educational Leadership*, *43*(3), 44–49.

Ponzio, R., Peterson, K., Miller, J., & Kinney, M. (1994). A Program Portfolio/Panel Review Evaluation of 4-H Sponsored Community-Based Social Action Projects for At-Risk Youth. *Journal of Educational Research and Development*, *28*(1), 55–65.

Popham, W.J. (1988). Judgment-Based Teacher Evaluation. In S. Stanley, & W. Popham (Eds.), *Teacher Evaluation: Six Prescriptions for Success* Arlington, VA: Association for Supervision and Curriculum Development, pp. 56–77.

Sanders, W.L., & Horn, S.P. (1995). The Tennessee Value-Added Assessment System (TVASS): Mixed Model Methodology in Educational Assessment. In A.J. Shinkfield, & D. Stufflebeam (Eds.) *Teacher Evaluation: Guide to Effective Practice*. Boston: Kluwer Academic Publishers, pp. 337–350.

Scriven, M. (1967). The Methodology of Evaluation. In R. Tyler, R. Gagne, & M. Scriven (Eds.), *AERA Monograph Review on Curriculum Evaluation: No. 1*. Chicago, IL: Rand McNally, pp. 39–83.

Scriven, M. (1973a). *Handbook for Model Training Program in Qualitative Educational Evaluation*. Berkeley, CA: University of California.

Scriven, M. (1973b). *The Evaluation of Educational Goals, Instructional Procedures and Outcomes*. ERIC Document Reproduction Service ED 079 394. Berkeley, CA: University of California.

Scriven, M. (1988). Duty-Based Teacher Evaluation. *Journal of Personnel Evaluation in Education*, *1*(4), 319–334.

Seldin, P. (1997). *The Teaching Portfolio: A Practical Guide to Improved Performance and Promotion/Tenure Decisions*, (2nd ed.). Bolton, MA: Anker Publishing Co.

Sergiovanni, T. (1977). Reforming Teacher Evaluation: Naturalistic Alternatives. *Educational Leadership*, *34*, 602–607.

Wheeler, P.H. (1994). Before you use Portfolios in Teacher Evaluation . . . Consider these Issues. *AASPA Report*, *1*(4), 8–9.

Wolf, K. (1991). The Schoolteacher's Portfolio: Issues in Design, Implementation, and Evaluation. *Phi Delta Kappan*, *73*, 129–136.

Wolf, K. (1996). Developing an Effective Teacher Portfolio. *Educational Leadership*, *53*(6), 34–37.

Wolf, K., Hagerty, P., & Whinery, B. (1995). Teaching Portfolios and Portfolio Conversations for Teachers and Teacher Educators. *Action in Teacher Education*, *17*, 30–39.

Wolf, K., Lichtenstein, G., Bartlett, E., & Hartman, D. (1996). Professional Development and Teaching Portfolios: The Douglas County Outstanding Teacher Program. *Journal of Personnel Evaluation in Education*, *10*, 279–286.

Wolf, K., Lichtenstein, G., & Stevenson, C. (1997). Portfolios in Teacher Evaluation. In J.H. Stronge (Ed.), *Evaluating Teaching: A Guide to Current Thinking and Best Practice*, pp. 193–214. Thousand Oaks, CA: Corwin Press.