



RELIABILITY OF PORTFOLIO : A CLOSER LOOK AT FINDINGS FROM RECENT PUBLICATIONS

PORTFOLYO DEĞERLENDİRMESİ : SON ÇALIŞMALARDAKİ BULGULARA YAKINDAN BİR BAKIŞ

Özge ÖZYALÇIN OSKAY*, Michael SCHALLIES**, İnci MORGİL***

ABSTRACT: In this review article, conventional portfolio assessment and new developments in portfolio assessment are investigated. The concept of portfolio, portfolio building steps, contents of portfolio, evaluation of portfolio, advantages, disadvantages and concerns in using portfolio as well as validity and reliability of portfolio assessment are discussed considering recent findings from research literature.

Keywords: portfolio, evaluation of portfolio, measurement and assessment methods.

ÖZET: Bu derleme makalede, geleneksel portfolio değerlendirmesi ve portfolio değerlendirmesindeki yeni gelişmeler değerlendirilmiştir. Portfolyo kavramı, portfolyo hazırlama basamakları, portfolyonun içindekiler, portfolyonun değerlendirilmesi, portfolyo kullanımının avantaj ve dezavantajları, portfolyo değerlendirmesinin güvenilirliği ve geçerliği literatürdeki en son bulgular ışığında değerlendirilmiştir.

Anahtar sözcükler: portfolyo, portfolyonun değerlendirilmesi, ölçme ve değerlendirme metodları

1. INTRODUCTION

Education in the late twentieth century has experienced a multitude change in trends, such as reexamination of the narrowness of long honored curricular emphases, instructional and learning paradigm shifts and experimentation with variations in classroom scheduling. There is a broad trend that influences curriculum content and methodology. This is a trend to focus on learners: - their characteristics, needs, interests, learning styles, etc - and to engage them in active construction of their knowledge, reflecting constructivist thinking, which could be dated back to John Dewey's Benchmark Theory of Learning by Doing. In the course of this development alternative methods to determine student learning came into being, supplementing the traditional methods. The word assessment has become commonplace among educators today as they struggle to keep pace with the ever-increasing complexity in their daily professional lives (Pennsylvania Assessment Through Themes Project 1999).

It is thought that one of the keys to effective student learning is self-assessment. This allows students to reflect, redirect and confirm their learning process (Pennsylvania Assessment Through Themes Project 1999). Portfolio is one of the self-assessment methods that may be used. Portfolio started to be used first in the '80s in order to enhance the quality of education (Haecker, 2002).

Paulson et al. (1991) have defined portfolio as a concept that could be used in various ways. However, according to Paulson et.al. a portfolio can be defined as a purposeful collection of student work that exhibits students' efforts, progress and achievements in one or more areas of learning. It must include student participation in selecting contents, criteria for selection, criteria for judging merit and evidence for student self-reflection (Paulson et. al. 1991).

Portfolios can be used in a variety of ways for student learning, e.g. for a specific course, for a prescribed block of courses or for a complete education program as an assessment organizer to demonstrate individual learner's knowledge, skills, and dispositions gained from selected learning activities. Also, portfolios can be used as a tool for teacher pre-service education, a tool for in-service

* PhD, Hacettepe University, ozge@hacettepe.edu.tr

** Prof. Dr., University of Education Heidelberg, schallies@ph-heidelberg.de

*** Prof.Dr., Hacettepe University, inci@hacettepe.edu.tr

training and a tool for interventions (Stader von Krosigh & Neely 2001). They can be presented in the traditional paper pencil design, in an electronic format or in some combination (Stader & Hill 2002).

Portfolios provide an essential link for students, teachers, and parents to the educational process. They are an effective tool for students and educators to identify, clarify and achieve goals that were not previously assessable. Thoughtfully implemented, portfolios can provide the bridge for students to become successful life-long learners (Schallies & Dumke, 2007). They are also a qualitative research instrument for gaining insight into individual learning processes in complex learning environments (Eysel, 2006).

A portfolio is student-centered and used to foster student self-evaluation. It provides greater depth and breadth of student work for assessment. Teachers may use it as a curriculum, assessment and professional development tool (Pennsylvania Assessment Through Themes Project 1999).

Also, portfolio assessment is an increasing common trend in teacher education programs because the nature of portfolios allows the demonstration of a variety of strengths and competencies essential to effective teaching (Gadbury-Amyot, 2003).

2. KEY STEPS IN BUILDING A PORTFOLIO

There is wide-spread agreement among practitioners that a portfolio should at least have the following building blocks:

Context definition: In this first step, the purpose of the portfolio, temporal frame, objectives and standards should be determined.

Collection: During the working phase at the concrete project, all documents that were found in the connection with the project goals should be gathered. These documents can be for example: Reflection documents, documents that are prepared during the work process (articles, instructions, own designs), protocols, video parts over a groups meeting etc. (Collins, 1992).

Selection: In the third step, an established selection of the documents takes place, that are important for the presentation of the results and the individual way from the viewpoint of the learners. Here, a reflexive process must take place.

Reflection: Beyond the presentation of examples of students' work, portfolios contain reflective statements written by the student. In these written reflections, students need to demonstrate to the reader the value of their work examples included in the portfolio and what the examples demonstrate about their intellectual growth. Self-reflection requires analysis and synthesis of thought and action, encouraging active involvement and a sense of ownership in the development of the portfolio and of one's own learning (Gadbury-Amyot et.al. 2003).

Reflection is a process involving meta-cognition that enables students to evaluate work based on established criteria. In this step students direct their own learning.

In this step, students raise queries in order to answer the question of how they learn and show it. In the step of the content selection, students prepare their portfolios by asking the following questions to themselves: What were the difficult points? What would I like to realize once more? What did I do best until now? What was the most surprising for me? Why did I choose this content for my portfolio? What does it mean to me? When I should prepare this portfolio once more would I prepare it like this again?. In this way students decide what, how and why to learn themselves and take the responsibility of their learning (Wellensiek; Lembens & Schallies 2001).

Projection: In this last step, the knowledge about students' own learning as brought out in the reflection is summarized, and appropriate measures are taken for further work (Haecker, 2003).

Contents of Portfolio

Contents of a portfolio may vary depending on the purpose of the portfolio. But generally a portfolio contains at least:

- Title Page
- Table of contents
- Segments.

Each thematic unit (essential question) may contain:

Introduction: Defines the content of a specific segment of the portfolio and the standards related to this segment.

Knowledge Base: Summary of the knowledge gained or acquired.

Application Base: How this knowledge can be applied.

Reflection: What the student thinks or believes about how the learning activities may relate to their own future practice. These reflections can be compared to the pre-reflection writing to determine and demonstrate growth (Stader & Hill, 2002).

3. EVALUATING AND ASSESSING PORTFOLIOS

Since portfolios do not fit the traditional system of lecture, test and grade format and contain personalized materials with few points of objectivity that allow comparison, it is difficult to assess them. Teachers have difficulties in grading and evaluating portfolios. Danielson and Abruthyn (1997) have suggested considering some general points for evaluation. These are:

- **Grades can be assigned only to items in assessment portfolios.**

Pieces in working portfolios should be evaluated against established and public criteria by both the teacher and the student. Those evaluations should be formative in nature only, however, and should serve as the basis for further instruction. Items in a display portfolio are placed there at the discretion of the student, and reflect the student's own judgment of high quality work. A teacher's grade, imposed over the student's own judgment, is superfluous.

- **Items can be evaluated by using a scoring guide or rubric which the students themselves helped to create:**

The principles at work in the evaluation of individual pieces are derived from the assessment of complex performance and involve clear and significant criteria, descriptions of levels of performance, and attention to reliability.

- **Clear guidelines for students about their work with the portfolio should be established:**

If students are to complete a reflection sheet for each item in the portfolio, provide instruction and modeling of the ways to accomplish such a task. If they are to engage in goal setting, make sure they know how that should be done. If guidelines have been given, if instruction has been provided, and if students have had the opportunity to acquire and practice the skills involved, evaluating their performance in these areas is justified. These skills, however, are new to most students and some opportunity to learn them will probably have to be provided, with formative assessment, before students are ready to be evaluated.

- **Clear guidelines for evaluating assessment portfolios as a whole for completeness and organization should be established:**

If a language arts portfolio is required to contain a persuasive essay, a descriptive essay, a narrative piece, a graphic organizer, an interview and essay, and a major report, that portfolio is incomplete if one or more of those pieces are missing or incomplete. Furthermore, an assessment portfolio should be comprehensible and organized so that a reader can navigate all the different sections. Reflection sheets should be attached to their documents, earlier drafts should be clearly labeled, and all pieces should be dated. If these aspects of the portfolio are to be evaluated, however, the criteria and performance levels should be clearly articulated in advance and students must have the opportunity to learn the necessary skills (Danielson & Abutryn, 1997).

In order to give students feedback about their work and tell them whether they address the portfolio dimension properly, grades can be given to the portfolios. For this reason rubrics can be used.

The term rubric has evolved into the term used to describe scoring guides for assessing the quality of student learning (Stader et.al. 2002). All rubrics have some commonalities, these are:

Evaluative criteria, quality definitions and scoring strategy (Popham, 1997; Goodrich 1997 & Andrade 2000).

All effective rubrics

- Contain three to five teachable evaluative criteria (Popham, 1997).
- Contain evaluative criteria that present key attributes of the skill being assessed (Popham, 1997).
- Avoid unclear and unnecessary negative language (Goodrich, 1997).
- List criteria and articulate levels of quality (Andrade, 2000).
- Involve students in the development, critique and use of evaluative criteria (Stader et.al. 2002).

The growth in portfolio assessment in the last few years is remarkable. Portfolios are starting to be used in all sorts of ways to assess learning in many subjects (Eysel, 2006).

Many researchers used different methods for the assessment of portfolio: Adamchik (1996) developed a subject specific portfolio model to demonstrate student goal setting, growth and achievement and reflection/staff assessment. In his study, he used the combination of assessment and conferencing. He used conferencing in order to communicate with the student concerning his/her portfolio. His idea was to have each student come in for a conference and defend his/her portfolio. During the conference they discuss the portfolio and they decide which grade should be given together.

Fenwick (1996) gave an example of a portfolio rubric for assessing portfolios documenting the growth of people enrolled in a certificate program in adult and continuing education. The portfolios collected items showing their growth of skill and understanding in program planning, learning and development, foundations of education and facilitating adult education.

In their study Wilcox and Tomei (1999) used portfolios as a professional development tool. They thought that portfolio is more than a place where teachers organize their teaching materials; it is a place where teachers record their thinking and learning about their teaching. They defined reading, writing, thinking, interacting and demonstrating as essential elements of a professional portfolio. And also they used rubrics for the assessment of the portfolio. While they organized their portfolio according to these five essentials so their rubric contains these five essential steps.

In his study Scott (2005) aimed to examine how reflective writing is solicited, taught, composed and assessed within the state writing portfolio curriculum in Kentucky. The Kentucky state assessment was designed to promote the use of portfolios in classes and provide information concerning students' learning for the state's system of teacher and school accountability. Each year every public school in Kentucky receives a numerical score based on various measurements of student performance including writing portfolios. This score is measured according to the accountability index and every 2 years, the averages of the scores are used to evaluate schools' abilities to progress toward the curriculum's long term numerical goals. All 12th grade writing portfolios are scored holistically according to a rubric developed by the state with each receiving a rank of Novice, Apprentice, Proficient and Distinguished. These scores become a part of the states system of accountability. Scott (2005) examined reflective letters from the portfolios of the 12th class students. He collected data through interviews, observations, and textual analysis. After the examination of the data he observed that although it causes a little bit anxiety for the students, with this kind of curriculum, students and the teachers exhibited a sophisticated understanding of portfolio pedagogy.

In their study Beck et.al (2005) compared the effects of different portfolio curricula (one summative accountability portfolio, two formative teacher development portfolios and one mixed formative/summative portfolio) on pre-service and beginning teachers' professional development. The sample of the study consisted of 207 teachers who were in credential or Master of Arts in teaching programs at a research university. As data collection tool he used the Electronic Portfolio Assessment Scale (ePAS). The ePAS consisted of 34 self-report statements, of which 17 were positively and 17 were negatively phrased statements about various effects of portfolio making such as self-discovery, increased understanding of reflective practice, greater knowledge of students, backwards planning,

better understanding the role of assessment in planning and instruction and support of the value of collaboration. The internal consistency of the scale was 0,92. After the evaluations no significant differences were found between the mean scores of male and female teachers concerning professional development on any of the portfolios or factors. This result suggested that electronic portfolios might provide a useful technique to enhance teacher development, irrespective of their sex, and that narrative methods were not unusually useful to females, as had been suspected.

In their study Gülbahar and Yilmaz (2006) used project based learning method in the Design, Development and Evaluation of Educational Software Course which is required for senior students of the Department of Computer Education and Instructional Technology, and which is part of the Faculty of Education at Başkent University (Ankara). In the study students were requested to prepare software for Instructional Technology and Material Development course with project based learning method. And for the evaluation of the application they used e-portfolios. The e-portfolios included written reports, multimedia presentations, statistical analyses and two versions of the software. During the study, semi-structured interviews were conducted with each student in order to gather students' perceptions about the project based learning approach. And for collecting data about overall satisfaction of the course and instructor, a likert-type questionnaire was used. After the evaluation of the interviews and the questionnaire, it was observed that students had high level of satisfaction about the course and instructor, and students agreed that project-based learning produced successful results for their own learning.

Schallies (2005) developed a Project in order to investigate the effects of portfolio on students' self-directed learning and life-long learning. This concept comes from "Professional Development Schools" approach. Portfolio is applied to classroom work and the effect of portfolio on students' self directed learning is investigated. At the end of the study the findings showed that portfolio is an effective tool for students' learning and for this study male students are more successful than female students.

The Postgraduate Certificate of Higher Education (PCHE) at the University of Sheffield in UK was established on the principle that the best way for students to learn about teaching is through enquiry into their own experiences. With this aim to support enquiry process, Harland (2005) involved portfolio as a component of a new teacher education curriculum. In his study he outlined a four-year case study between 1996 and 2000, which described the way a portfolio was introduced, evaluated and developed on a pre-service teacher education course for future university lecturers. The portfolio aimed to provide "authentic enquiry" that focused on students' self determination of the process, rather than the outcomes of the learning. Initially, portfolios were evaluating formatively during supervisory meetings and each student decided what part of their portfolio should remain private and what the tutor might read and comment on. In the second phase development, formative judgments of the work were no longer made and portfolios became private documents. For data collection, written student evaluations were produced annually during these four years; students were asked to write about their portfolio experiences and learning. A student survey was administered in 1996 to ascertain the costs and benefits of portfolio as a private or public document. Online asynchronous discussions with students and teachers during 1997 focused on portfolio experiences. All contributions were printed and archived.

All stages of the applications and evaluations showed that the portfolio helped the majority to explore their practice and gain insight into themselves as teachers and learners. Some students found reflective writing tasks difficult, they found enough satisfaction and challenge in the course activities. This unique experience helped student teachers understand their learning and teaching practice, and for many it became a challenge that lived up to a concept of authentic enquiry.

4. VALIDITY AND RELIABILITY OF PORTFOLIO ASSESSMENT

Because of being different from the traditional teaching and assessment methods, application and usage of portfolios, educators and students are often sceptical about the whole idea and

preparation of portfolios (Scheurman, 1988). For researchers the overarching question remains whether portfolio assessment is psychometrically sound and can produce psychometrically defensible results.

The Standards for Educational and Psychological Testing (1999) define validity as “the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests”. This definition considers validity as a unified concept and is based in part on Messick’s theory of validity (Messick, 1989; Miller, 2000). Messick argued that validity is a unitary concept and proposed that validation of performance assessment, as with all other forms of assessment, should incorporate six aspects of construct validity: content, substantive, structural, generalizability, external, and consequential. Reliability has been defined as the measure of the degree of consistency in examinee scores over replications of a measurement procedure (Brennan 2001).

Gadbury-Amyot et.al (2003) examined the validity and reliability of portfolio assessment using Messick’s unified framework of construct validity. Theoretical and empirical evidence was sought for six aspects of construct validity. Seven faculty raters evaluated twenty student portfolios using a primary trait analysis scoring rubric. The relationship between portfolios and the Central Regional Dental Testing Service (CRDTS) examination was both weak and nonsignificant ($r = .19$; $p > .05$). A fully crossed, two facet generalizability (G) study design was used to examine reliability. ANOVA demonstrated that the greatest source of variance was the scoring rubric itself, accounting for 78 percent of the total variance. The smallest source of variance was the interaction between portfolio and rubric (1.15 percent). Faculty rater variance accounted for only 1.28 percent of total variance. A phicoefficient of .86, analogous to a reliability coefficient in classical test theory, was obtained in the decision study by increasing the subscales to fourteen and decreasing faculty raters to three. In conclusion, the pattern of findings from this study suggests that portfolios can serve as a valid and reliable measure for assessing student competency.

Driessen et.al (2006) investigated the validity of portfolio assessment and the qualities that determine the ratings. In their study they developed an instrument. The Portfolio Analysis Scoring Inventory, to examine the impact of form and content aspects on portfolio assessment. The Inventory consists of 15 items derived from interviews with experienced mentors, the literature and the criteria for reflective competence use in the regular portfolio assessment procedure. Forty portfolios, selected from 231 portfolios for which ratings from the regular assessment procedure were available, were rated independently by 2 researchers, using the inventory. Regression analysis was used to estimate the correlation between the ratings from the regular assessment and those resulting from the Inventory items. The results show that the ratings were primarily determined by quality of reflection. Quality of reflection was the only item to make a significant contribution (64%) to the explanation of the variance of the regular portfolio ratings. The non-significant effects found for lay-out, spelling and grammar and structure as compared with quality of reflection may be attributable to the mentoring and assessment training, which thus appears to positively affect the validity of the assessment procedure.

Koretz, Stecker and Deibert (1993) reported disappointing results from a Vermont’s statewide portfolio assessment project. Students in grades four and eight developed mathematics and writing portfolios that were rated by classroom teachers. After the analysis inter-rater reliabilities ranging from 0.28 to 0.60, a level of agreement considered not sufficient enough (Gadbury-Amyot, 2003).

Derham and Diperna (2007) used the latest version of portfolios, digital portfolios. The purpose of the study was to examine initial evidence regarding the reliability and validity of a digital professional portfolio for assessing the instructional competencies of preservice teachers. Pearson product-moment correlation coefficients, Cronbach's alpha, and Cohen's kappa coefficient. Evidence supporting the use of the DPP was found through theoretically acceptable relationships with other teacher assessments. Weaknesses of the DPP also were identified, including low percentage of exact agreement between scorers.

Gelinas (1998) in her dissertation investigated the reliability and validity of portfolio scores in evaluating the teaching competence of preservice foreign language teachers. It was determined that the

portfolios are a valid measure of students' pedagogical content knowledge (ability to present content in a learnable format) based on results from correlational and regression analyses. It was also indicated that the portfolio scores were reliable and accurate measure of students' competence. The inter-rater agreement level was high ($r=.81$) as was the score reliability index ($r=.90$) in the Rasch-model FACETS analysis.

In their study, Rees and Sheard (2004) aimed to determine the reliability of assessment criteria used for a portfolio at the University of Nottingham. Two independent analysts assessed a random sample of portfolios using criterion-referenced assessment. Students' performances were examined against subjective items in five areas: 1) portfolio structure, 2) level of critical reflection, 3) level of skills development, 4) use of documentary evidence, and 5) use of relevant literature. These subjective judgments were later converted into quantitative scales ranging from 0 to 3 so that interrater reliability could be established. The level of agreement between the two analysts for the total percentage score was established using an intraclass correlation coefficient and for the individual items using weighted kappa coefficients. The level of agreement between the two raters for the total percentage score was 0.771 (95% CI $k= 0.678, 0.840$), as measured by an intraclass correlation coefficient. The levels of agreement between the two raters for the individual items of the assessment criteria ranged from $k= 0.359$ to $k= 0.693$. This study provides some support for the summative assessment of portfolios. The findings suggest that discussion and negotiation between independent assessors can enhance the reliability of assessment criteria (communicative validation). Therefore, medical educators are encouraged to use such procedures in the summative assessment of portfolios.

However, these studies suggest that there are some general principles that can be used to guide large-scale portfolio-based assessment. To achieve high levels of interrater reliability of around 8-8,9 it appears that portfolios should be carefully introduced to well prepared students and should be of uniform content. They should be marked by experienced, trained scorers, who use clearly articulated criteria, have a shared understanding of the purpose of assessment and a deep understanding of expected student performance.

According to Stiggins (1987) the reliability is maximized by using clear scoring criteria, training raters, gathering enough samples of performance and providing for standard assessment conditions (Gadbury-Amyot, 2003).

5. RESULTS

From these recent findings it can be concluded that portfolio assessment is an appropriate and reliable assessment method for individual learning, giving more insight into the process of learning rather than being a tool for product assessment. It balances assessment and learning through emphasis on self reflection and meta-cognitive approaches to teaching and learning. It gives the teacher the opportunity to synthesize thinking and learning in a creative way, while illustrating engagement and reflection, deep understanding of content, meaningful connections and skills in oral communication. The rubrics can serve as tools to guide the portfolio construction, while adding dimension to the portfolio assessment process. Portfolio is a common assessment method, which can be use in different educational levels in different subjects. For this reason extra importance can be given to the portfolio assessment in teacher education programs of the universities.

REFERENCES

- Adamchik, C. (1996). The design and assessment of chemistry portfolios. *Journal of Chemical Education*, 73(6), 528-31.
- American Educational Research Association, American Psychological Association, and National council on measurement in education. *Standards for Educational and Psychological testing* (1999). Washington, DC: American Educational Research Association.
- Andrade, H. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57(5), 13-18.
- Beck, R. J., Livne, N. L. & Bear, S. (2005). Teachers' self-assessment of the effects of formative and summative electronic portfolios on professional development. *European Journal of Teacher Education*, 28(3), 221-244.

- Brennan, R. (2001). An essay on the history and future of reliability from the perspectives of replications. *Journal of Educational Measurement*, 38, 295-317.
- Collins, A. (1992). Portfolio for science education: Issues in purpose, structure and authenticity, *Science Education*, 76(4), 451-463.
- Danielson, C & Abrutyn, L. (1997). *An introduction to using portfolios in the classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Derham, C., Diperna, J., (2007). Digital Professional Portfolios of Preservice Teaching: An Initial Study of Score Reliability and Validity. *Jl. of Technology and Teacher Education*, 15(3), 363-381.
- Driessen, E. W., Overeem, K., Tartwijk, J., Van der Vleuten, C & Muijtjens, A. M. M., (2006). Validity of portfolio assessment: Which qualities determine ratings? *Medical Education*, 40, 862-866.
- Eysel, C. (2006). *Interdisziplinäres Lehren und Lernen in der Lehrerbildung. Eine empirische Studie zum Kompetenzerwerb in einer komplexen Lernumgebung*. Berlin: Logos Verlag.
- Fenwick, T. (1996). *Assessing adult learning using portfolios*, Paper presented at the Atlantic Teaching Showcase Conference. Université St. Anne, Church Point, Nova Scotia, October 25, 1996, Canada.
- Gadbury-Amyot, C. (2003). *Validity and reliability of portfolio assessment of competency in a baccalaureate dental hygiene program*, A Dissertation in Education, Doctor of Education, University of Missouri-Kansas City.
- Gadbury-Amyot, C, C., Kim, J., Palm R. L., Mills, E., Noble, E & Overman, P. (2003). Validity and reliability of portfolio assessment of competency in a baccalaureate dental hygiene program, *Journal of Dental Education*, 67(9) 991-1002.
- Gelinas, A. (1998). *Issues of Reliability and Validity in Using Portfolio Assessment to Measure Foreign Language Teacher Performance*, Dissertation, Doctor of Philosophy in the Graduate School of the Ohio State University.
- Goodrich, H. (1997). Understanding rubrics. *Educational Leadership*, 54, 14-17.
- Gülbahar, Y. and Tinmaz, H. (2006). Implementing Project-Based Learning and E-Portfolio Assessment in an Undergraduate Course, *Journal of Research on Technology in Education*, 38(3), 309-327.
- Häcker, T. (2002). Der Portfolioansatz - die Wiederentdeckung des Lernsubjekts? *Die Deutsche Schule*, 94(2), 204-216.
- Häcker, T. (2003). Wissen und Handeln reflektieren: Die Qualität des Projektlernens mit Portfolios entwickeln: In: Eyerer, P. and Krause, D. (Eds.), *Bildung ist Wissen und Handeln*, 6. TheoPrax Tagung, 2003 Sep 25; Pfnitzal: Fraunhofer Institut für Chemische Technologie (ICT). (pp. 17-31).
- Harland, T. (2005). Developing a portfolio to promote authentic enquiry in teacher education, *Teaching in Higher Education*, 10(3), 327-337.
- Messick S. (1989). Validity in performance assessments. In: Linn RL, (Eds.) *Educational Measurement*. New York: American Council on Education and Macmillan, 13-104.
- Miller, M.D & Linn, R. L., (2000). Validation of performance-based assessment. *Appl. Psychol. Meas.*, 24, 367-78.
- Paulson, F. L., Paulson, P. R. and Meyer, C. A. (1991). What makes a portfolio a portfolio? *Educational Leadership*, February, 60-63.
- Pennsylvania Assessment through Themes Project, (1999). *The literacy dictionary. The scass science portfolio implementation guide*. Retrieved in December 2006 from www.pde.state.pa.us/fam_consumer/lib/fam_consumer/20/23/portig.pdf.
- Popham, W. J. (1997). What's wrong- and what's right- with rubrics. *Educational Leadership*, 55, 72-75.
- Rees, C. E., & Sheard, E. (2004). The reliability of assessment criteria for undergraduate medical students' communication skills portfolios: the Nottingham experience, *Medical Education*, 38(2), 138-144.
- Schallies, M., (2005). Die Auswirkungen von Portfolioarbeit in der Schule auf Eigenverantwortung und Selbststeuerung der Schüler, *H. U. Journal of Education*, 29, 167-176.
- Schallies, M., & Dumke, J. (2007). *Lebenslanges Lernen. Ein forschungs- und Entwicklungsprojekt zur Stärkung von Eigenverantwortung und Selbststeuerung im Rahmen schulischen Lernens*. Heidelberg: Mattes Verlag.
- Scheurman, G. (1998). From behaviorist to constructivist teaching, *Social Education*, 69(1), 6-9.
- Scott, T. (2005). Creating the subject of portfolios. *Written Communication*, 22(1), 3-35.
- Stader, D. L., & Hill-Winstead, F. (2002). *Portfolios, performance assessments, and standards based learning in educational leadership*, Paper presented at the annual National Council of Professors of Educational Administration Conference, Burlington, Vermont. ERIC Document Reproduction Services ED 480133.
- Stader, D., Von Krosigh, K., & Neely, R. (2001). The professional portfolio for aspiring principals. *The AASA Professor*, 24(4), 15-19.
- Wellensiek, A., Lembens, A., & Schallies, M. (2001). Lernen mit dem Portfolio, *Unterricht Arbeit+Technik*, 9, 21-23.
- Wilcox, B., & Tomei, L., A., (1999). Assessing teacher performance with a portfolio rubric, *English Leadership Quarterly* 22(2), 6-10.
- Yılmaz, G ve Tinmaz, H. (2006). Implementing project based learning and e-portfolio assessment in an undergraduate course, *Journal of Research on Technology Education*, 38(3), 309-327.

GENİŞLETİLMİŞ ÖZET

Eğitim alanında yaşanan değişim ve gelişimlerle birlikte günümüzde eğitimciler arasında değerlendirme kelimesi çok yaygın olan bir kelimedir.

Etkili öğrenmenin anahtarlarından birinin kendi kendini değerlendirme olduğu düşünülmektedir (Pennsylvania Assessment Through Themes Project 1999). Portfolyo da etkin olarak kullanılabilir kendi kendini değerlendirme metotlarından biridir. Paulson vd. göre portfolyo, bir ya da daha fazla öğrenme alanında öğrencinin çabalarını, çalışmasını ve başarısını gösteren çalışmalarının topluluğudur. Portfolyo, içeriğin seçiminde öğrenci katılımını, seçim ve değerlendirme için bir kriteri ve öğrencinin kendi kendini değerlendirmesini içermelidir (Paulson et. al. 1991).

Portfolyolar, özel bir ders için, blok derslerde ya da tam bir eğitim programında öğrencilerin bilgilerini, yeteneklerini ve seçilmiş bir öğrenim aktivitesinden kazanımlarını göstermek amacıyla pek çok şekilde kullanılabilir. Ayrıca, portfolyolar öğretmen eğitiminde servis öncesi eğitimde bir araç olarak da kullanılabilir (Stader von Krosigh and Neely 2001). Değerlendirme için öğrenci çalışmalarına büyük bir derinlik katar (Pennsylvania Assessment Through Themes Project 1999).

Uygulamacılar portfolyonun en azından aşağıdaki alt başlıklara sahip olması gerektiği konusunda hem fikirdirler:

İçeriğin tanımlanması

Toplama

Seçme

Yansıtma

Sunum (Haecker 2003).

Portfolyonun içeriği amacına göre değişiklik gösterebilir. Ancak genel olarak bir portfolyo en azından şu başlıkları içermelidir.

- Başlık Sayfası
- İçindekiler
- Bölümler:

Her bir ünitenin ya da (özel bir sorunun) içerebileceği bölümler:

Giriş

Bilgi Bölümü

Uygulama Bölümü

Yansıtma (Stader & Hill, 2002).

Portfolyolar geleneksel ders sistemine uymadıklarından değerlendirilmeleri zordur. Öğretmenler portfolyoların değerlendirilmesinde ve puanlanmasında zorluklar yaşamaktadırlar. Danielson ve Abruthyn (1997) değerlendirme için bazı genel noktalar önermiştir. Bunlar:

- Portfolyolar değerlendirilirken her bir bölüme ya da her bir maddeye puanlar verilebilir.
- Maddeler ya da bölümler, oluşturulmasına öğrencilerin yardım ettiği bir değerlendirme ölçeği ya da puan skalası yardımıyla değerlendirilebilirler.
- Öğrencilere portfolyo ile yapacakları çalışmaları için açık kılavuzlar sunulabilir.
- Portfolyonun bütünlüğü ve organizasyonu için açık bir kılavuz hazırlanmalıdır.

Öğrencilere kendi çalışmalarını için geri dönüt sağlamak ve portfolyolarını doğru hazırlayıp hazırlamadıkları hakkında bilgi vermek amacıyla portfolyolar puanlanabilir. Bu amaçla da çeşitli değerlendirme ölçekleri kullanılabilir (Stader et.al. 2002).

Bütün ölçeklerin bazı ortak noktaları vardır. Bunlar: Değerlendirme kriterleri, kalite tanımlamaları ve puanlama stratejileridir (Popham, 1997; Goodrich, 1997 & Andrade, 2000).

Pek çok araştırmacı portfolyonun değerlendirilmesi için farklı metodlar kullanmışlardır. Adamchik (1996) öğrencilerin hedeflerini belirlemelerini ve göstermelerini, portfolyoyu geliştirmedeki başarılarını, yansıtma ve takım değerlendirmelerini ölçmek amacıyla konuya özel portfolyo modeli geliştirmiştir. Çalışmasında değerlendirme ve konferans yöntemini beraber kullanmıştır. Öğrencilerle kendi portfolyoları hakkında iletişim kurmak amacıyla konferans yöntemini kullanmıştır. Öğrencileri kendi portfolyolarını savunmak ve portfolyolarının eksik ve güzel yanlarını tartışmak amacıyla onlarla konferans şeklinde bir iletişime girmiştir. Konferans boyunca her öğrenci ile kendi portfolyosu tartışılmış ve hangi notun verileceğine beraber karar verilmiştir.

Fenwick (1996) bir sertifika programına devam eden yetişkinlerin gelişimlerini değerlendirmek için bir portfolyo ölçeği geliştirmiştir. Portfolyolar öğrencilerin yeteneklerinin gelişimi, program planlamayı anlamaları, öğrenme ve gelişme eğitiminin temelleri ve yetişkin eğitiminin geliştirilmesi konuları ile ilgili bölümleri içermektedir.

Wilcox and Tomei (1999) ise çalışmalarında portfolyoyu bir mesleki gelişim aracı olarak kullanmışlardır. Araştırmacılar portfolyonun öğretmenlerin öğretim materyallerini organize ettikleri bir dosyadan çok daha fazla şey ifade ettiğini, öğretmenlerin düşüncelerini not alabilecekleri, kendi öğretimleri hakkında bilgi edinebilecekleri bir araç olduğunu düşünmektedirler. Bir mesleki portfolyonun temel elementlerini ise okuma, yazma, düşünme, iletişim kurma ve sunma olarak belirlemişlerdir. Ayrıca portfolyonun değerlendirmesi için ölçekler kullanmışlardır. Kullanılan ölçeklerde ise belirtilen temel elementler bulunmaktadır.

Geleneksel öğretim ve değerlendirme metodlarından farklı olmasından dolayı öğrenci ve eğitimciler genellikle portfolyoların uygulanması ve kullanılması konusunda şüphecidirler. Araştırmacılar için kapsayıcı soru: portfolyo değerlendirmesinin psikometrik olarak geçerli ve güvenilir olup olmadığı ve psikometrik olarak savunulabilir sonuçlar doğurup doğuramayacağıdır. Gadbury-Amyot et.al (2003) portfolyo değerlendirmesinin geçerliğini ve güvenilirliğini Messick'in yapı geçerliğini kullanarak saptamıştır. Yapı geçerliğinin altı alt boyutu için teorik ve empirik kanıtlar araştırılmıştır. Üniversiteden yedi değerlendirmeci yirmi öğrenci portfolyosunu puanlama ölçeği kullanarak değerlendirmiştir. Sonuç olarak bu çalışmanın sonuçları portfolyoların öğrencilerin yeteneklerinin ölçülmesinde geçerli ve güvenilir bir ölçüm sağladığını öne sürmektedir. Driessen vd. (2006) ise çalışmalarında portfolyoların geçerliğini ve puanlama için gerekli özellikleri araştırmışlardır. Portfolyo değerlendirmesine şekil ve içerik boyutunun etkisini araştırmak amacıyla "Portfolyo Analiz Puanlama Envanteri" adı ile bir ölçek geliştirmişlerdir. Sonuçlar, puanlamanın yansıtmanın niteliği ile belirlendiğini göstermektedir. Portfolyo değerlendirmesinin ve puanlanmasının varyansının açıklanmasında yansıtmanın kalitesi en anlamlı katkıyı (64%) yapan madde olarak tespit edilmiştir. Her ne kadar bu çalışmalar büyük ölçekli portfolyo tabanlı değerlendirmeleri yönlendirmek için bazı genel prensipler önerse de 8-8,9 seviyesinde güvenilirlik elde edebilmek için portfolyolar dikkatlice iyi yetişmiş öğrencilere sunulmalı ve içerik açısından tek biçimli olmalıdır. Ayrıca, deneyimli, eğitimli, puanlama kriteri kullanan, portfolyonun amacını, değerlendirmenin amacını anlayan puanlamacılar tarafından değerlendirilmelidirler. Yapılan son çalışmalardan portfolyonun öğrencilerin öğrenmeleri için uygun ve güvenilir bir araç olduğunu göstermektedir.