Teachers of English to Speakers of Other Languages, Inc. (TESOL)

Portfolios, Power, and Ethics
Author(s): Brian Lynch and Peter Shaw
Source: *TESOL Quarterly*, Vol. 39, No. 2 (Jun., 2005), pp. 263-297
Published by: Teachers of English to Speakers of Other Languages, Inc. (TESOL)
Stable URL: http://www.jstor.org/stable/3588311
Accessed: 02/12/2014 11:12

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Teachers of English to Speakers of Other Languages, Inc. (TESOL) is collaborating with JSTOR to digitize, preserve and extend access to *TESOL Quarterly*.

http://www.jstor.org

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

Portfolios, Power, and Ethics

BRIAN LYNCH
Portland State University
Portland, Oregon, United States

PETER SHAW
Monterrey Institute for International Studies
Monterrey, California, United States

Portfolios have been used in a variety of ways for assessing student work. In education, generally, and more specifically in second language education, portfolios have been associated with alternative assessment (Darling-Hammond, 1994; Hamayan, 1995; Shohamy, 1996; Wolf, Bixby, Glenn, & Gardener, 1991). This article defines *alternative assessment* as representing a paradigm and culture that is different from traditional testing, requiring a different approach to addressing the issues of validity and ethics. We present a framework that integrates a consideration of how power relations determine the ethics and validity of assessment inferences. We then apply that framework to the assessment of student portfolios in a master of arts in TESOL (MA TESOL) program.

BACKGROUND

This article represents the confluence of two interests: the special problems of ethics and validity posed by portfolio assessment and the analysis of findings from a longitudinal study of portfolio assessment in an MA TESOL program. The literature reveals a degree of controversy and confusion concerning the use of portfolio assessment as an alternative to traditional testing. In this article, we show that portfolio assessment is more than merely one of many homogeneous *alternatives in assessment* (Brown & Hudson, 1998). In doing so, we argue that the portfolio, as an exemplar of alternative assessment, represents a different paradigm or culture that requires an approach to validity evidence (i.e., to establishing the trustworthiness of the inferences made from the assessment process) differing in certain critical aspects from the approach used in traditional testing. As we clarify this difference, we do not argue that one approach is better than the other. We also do not argue that alternative assessment is somehow automatically valid. Alternative

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

assessment, like traditional testing, must make a case for the validity of the claims made from the assessment process. Because alternative assessment represents a different research paradigm, a different cultural view about assessment, it requires a framework for making the case for validity that differs from the traditional testing approach (i.e., beginning with the establishment of *reliability*, or consistency of measurement, and then moving to other, statistical evidence that the measurement is capturing what it claims to be assessing). The framework required for alternative assessment must include the relations of power that exist in the assessment process.

In this article, we present key issues in assessment ethics and validity including considerations of power, describe the Monterey Institute of International Studies (MIIS) master of arts portfolio, outline an ethical-validity framework, then examine the portfolio in relation to that framework. The conclusion highlights the need for further investigation and thought.

It is important at the outset to issue the following disclaimer: Although the MIIS portfolio concept and process was developed and is conducted collaboratively, each participant (faculty member or student) still retains his or her own views on both the broad scope and the details. The account presented here, therefore, does not necessarily represent the views of all participants.

KEY ISSUES

When the MIIS MA TESOL program moved to portfolio assessment as its exit mechanism, it moved away from an existing, more traditional procedure: the closed-book comprehensive written examination. This move embraces the construct of alternative assessment.

Alternative Assessment

In its simplest form, alternative assessment is a turn to something new, a search for a wider palette of choices. This vision of alternative assessment appears in such characterizations as Shohamy's (1996) "portfolios, observations, peer assessment, interviews, simulations, . . . self assessment" (p. 153) or Darling-Hammond's (1994) "oral presentations, debates, . . . exhibitions, . . . videotapes of performances and other learning occasions, constructions and models, . . . experiments" (pp. 5–6). However, such listings of various learning activities and products converted into components of assessment will not, in the end, be satisfactory. Indeed, Brown and Hudson (1998) and Norris, Brown,

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

Hudson, and Yoshioka (1998, p. 3) argue that these are not new forms of assessment at all, that they do not represent an alternative in the sense of being completely different from the range of assessment practices that have gone before, and they suggest using instead the term *alternatives in assessment.* Perhaps, then, the true soul of assessment lies not in the components or tools, but in the perspective or set of assumptions motivating their use. Wolf et al. (1991) have characterized this notion as *assessment culture.* This culture assumes that assessment involves an investigation of developmental sequences in student learning, a sampling of genuine performances that reveal the underlying thinking processes, and the provision of directions and opportunities for further learning. Assessment culture also assumes that teaching, learning, and assessment practices are inseparable and integral; students should be active participants in the process of developing assessment procedures, including the criteria and standards by which performances are judged; both the process and product of assessment tasks should be evaluated; and the assessment results should be reported as a qualitative profile rather than a single score or other quantification (Birenbaum, 1996). In short, this characterization of alternative assessment represents a significant change from the traditional testing perspective. We further develop the contrast between the two approaches later in this article.

Although portfolios are probably the most commonly cited example of alternative assessment (see, e.g., Darling-Hammond, 1994; Hamayan, 1995; Shohamy, 1996; Wolf et al., 1991), not all portfolios are good exemplars of alternative assessment culture (Hamp-Lyons & Condon, 2000). For portfolios to be considered alternative assessment, the process of selecting and assembling components, the nature of the final product, and the reading, feedback, and evaluating procedures must demonstrate these features:

- The students actively participate in the selection of the portfolio components.
- The students reflect on this selection process, and their reflection is included in the portfolio.
- The process of creating and selecting the portfolio components is included in the evaluation.
- The evaluation contains elements of peer and self-assessment.
- The portfolios are evaluated by persons familiar with the individual students and their learning context.
- The students participate in deciding the criteria for evaluating the portfolios.
- The evaluation is reported qualitatively, as a profile or other detailed description of what the student has achieved.

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

Another way of thinking about portfolios as alternative assessment comes from Hamayan's (1995) distinction between "activities that yield alternative assessment information" and "ways of recording alternative assessment" (pp. 217–218). Here, *recording* refers to the process of making judgments concerning the information presented in the portfolio and reporting those judgments. *Alternative assessment information* results from portfolios that simply display samples of writing and other student products. *Alternative assessment recording* results from involving students in deciding which products to include and the basis for judging those samples, and including in the reporting format a qualitative component that considers process as well as products. It is possible, then, to have alternative assessment information without alternative assessment recording. Using alternative methods for both gathering and recording information results in *portfolio-based alternative assessment* (Wolf et al., 1991, p. 57).

Validity and Ethics in Assessment

Our examination of portfolios as an assessment procedure focuses on validity and ethics and how they are affected by power relations. By *validity* we mean the degree to which a portfolio assessment is accomplishing what it claims or intends to accomplish. Another way of stating this is the degree to which the inferences or conclusions drawn from the portfolio assessment process are believable or trustworthy. In making this determination, however, one has competing perspectives from which to choose, among them the traditional, *positivist* perspective and the alternative, *interpretivist* perspective.

We acknowledge the danger in reducing the variety of research perspectives to a simple dichotomy, positivist versus interpretivist. However, despite the diversity and, at times, overlap between *positivism, postpositivism, constructivism, hermeneutics,* and *critical theory* (see, e.g., Kincheloe & McLaren, 2000; Lynch, 2001; Pennycook, 2001; Phillips, 1990; Schwandt, 2000), certain distinctions can be made. The (arguably) dominant paradigm in TESOL, which we call *positivism* (knowing that it is no longer an active school of philosophy, but also recognizing that certain key tenets survive), differs in important ways from other paradigms that we group under the label *interpretivism.* We refer to *traditional testing* to indicate assessment practices that follow the positivist perspective, and *alternative assessment* to indicate practices that follow the interpretivist perspective.

One distinguishing characteristic is the way each approach conceptualizes the thing it is trying to know or assess. The positivist perspective involves seeing that thing, in this case language ability, as parallel to

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

objects and phenomena of the physical world; the interpretivist perspective views language ability as part of the social world, which is seen as essentially different from the physical world. By acknowledging this distinction, we are claiming that knowledge about language can be legitimately pursued in ways that do not follow the scientific method used in the physical and natural sciences.

Another characteristic that distinguishes these two paradigmatic perspectives has to do with how the inquirer (researcher, teacher, assessor) establishes the relationship with what he or she is trying to understand. The positivist perspective uses *objectivity* as a guiding principle, with the inquirer (in this case, the persons doing the assessment) required to stay neutral and disinterested in relation to the object of inquiry (in this case, teacher development in an MA TESOL program). In establishing this relationship, the inquirer claims a meaningful and important distinction between *facts*—the things he or she is attempting to know—and *values*—subjective biases that can potentially distort the ability to capture the facts. In addition to objectivity as a guiding principle, the positivist perspective is primarily interested in the discovery of causal relationships, which are seen as key to knowledge and to the goal of generalizing from the particular results of inquiry to larger populations of people (for a sympathetic treatment of this perspective, see Phillips, 1990).

The interpretivist perspective finds it impossible to separate facts from values and accepts the inherent subjectivity in any research conducted in relation to people, to the social world. In this view, the knowledge that the inquirer seeks, the abilities that he or she is trying to assess, is seen as socially constructed rather than as part of an independently existing reality. Because of this assumption, the notion of causality plays a different and less central role: Causal relationships are simply one possible construction or explanation for certain aspects of the social world that the inquirer is investigating. They are not taken to be universal laws that govern people and their actions, including the acquisition and use of language. The interpretivist views relationships not in terms of a unidirectional and fixed causality—one variable preceding and causing another—but as more complex and fluid, with mutual and shifting directions of influence. Most important, the interpretivist does not see relationships involved with aspects of the social world, such as language or other abilities, as external and independent of his or her attempts to understand them. Instead of attempting to achieve a true match between inquiry and reality, the interpretivist perspective understands reality as being constructed in and through the inquiry process (for a discussion of the various strands of interpretivism, see Schwandt, 2000).

Another issue central to this article is *ethics*. We use this term to refer to the notion of morality, or what constitutes good and proper behavior.

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

As with validity, ethical determinations can be made from obviously different perspectives. We formulate our determination around relatively basic, noncontroversial human rights, such as the right not to be deceived or physically, emotionally, or psychologically abused. However, we also explore the relationship between ethics and relations of power, as discussed by Foucault (1997a–1997e).

Validity From the Alternative Assessment Perspective

Traditional testing establishes validity by adhering to certain principles and gathering particular types of evidence. One principle is that a test must be reliable (i.e., it must provide scores that are consistent, usually estimated by determining how consistent the answers on the individual test items are across a representative group of test takers) before it can claim to produce valid inferences or conclusions. Moss (1994, 1996) challenges this tenet. In articulating an approach to validity from the alternative assessment perspective, Moss (1994) points out that making *reliability* the necessary (but insufficient) criterion for validity can potentially limit educators' ability to pursue innovation in educational practice. Instead of the traditional approach to assessment, which requires objectivity and a supposedly unbiased, external person or persons making the judgments or a test instrument that removes as much subjectivity as possible (and provides a single score for each test taker), Moss argues for a *hermeneutic* approach. This approach represents an interpretivist research perspective and responds to some of the essential qualities of alternative assessment outlined earlier: It acknowledges the importance of the context of assessment and begins to formulate validity as a consensus reached through dialogue. This dialogue happens between stakeholders in the assessment—teachers, students, parents—not between disinterested external experts. Moss then poses a controversial question—can there be validity without reliability?—and provides an answer: "When reliability is defined as consistency among *independent* measures intended as *interchangeable,* the answer is, yes" (1994, p. 10, emphasis in original).

Moss (1996) also challenges the principle in traditional testing and assessment that the goal is generalizability. This goal means that educators want to be able to generalize from one performance, or from one set of test answers, to all such performances, contexts, or areas of knowledge. Validity, from this perspective, is the degree to which the specific performance or test score matches constructs or meanings that educators have predetermined as the object of assessment. Moss argues that this approach to validity will not allow educators to investigate and assess the sorts of social phenomena that they are interested in. The meanings

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

and interpretations that educators want cannot be predetermined and can only be understood and assessed within their particular contexts. Rather than sounding the death knell for generalization, it can be approached from a different perspective: A generalized understanding of some particular ability or knowledge can be advanced from a specific assessment context, and in hermeneutical style, this tentative or preliminary generalization can be compared with future contextualized assessments. Comparing current generalized understandings with new, contextualized interpretations is done in a reciprocal fashion rather than in the scientific hypothesis-testing mode; that is, there is a give and take between one interpretation and the next that results in a reshaping or reconceptualization of both (for a discussion of this approach to language test validity, see Hamp-Lyons & Lynch, 1998).

The important contrasts with traditional validity raised by alternative assessment, then, include challenges to the notions that reliability—internally consistent test scores or agreement between raters of language performance samples—is prerequisite for validity and that generalizability is the ultimate goal. Alternative assessment suggests that disagreement between raters in an assessment context may be a source of important, valid information, and that validity involves reaching a consensus among people who are closely involved with that context. Within this perspective on validity, the "portfolio [is] a vehicle connecting student to interaction with the curriculum rather than a product of the curriculum" (Wile & Tierney, 1996, p. 213).

The Relationship of Ethics to Validity

As stated earlier, we view ethical practice as protecting research participants' basic right not to be harmed (socially, psychologically, emotionally, physically) and not to be coerced or manipulated. Hamp-Lyons (1989) points out that ethics is properly seen as an essential aspect of validity. This interrelationship can be thought of in terms of the following questions:

- What should be the relationship between assessor and the person being assessed?
- What should the person being assessed be asked to do to demonstrate his or her ability?
- What does the assessor believe the nature of that ability to be?
- How does the assessor decide or, perhaps, who decides what counts as evidence?

The way these questions are answered will differ, depending on the assessor's perspective (i.e., traditional testing or alternative assessment).

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

It is arguable that the basic ethical questions should remain the same, regardless of the approach to assessment; that is, whether traditional or alternative assessment is being carried out, test-takers' rights not to be harmed or coerced remain constant. However, depending on the validity perspective, how educators define *harm* and *coercion* and the degree of ethical responsibility that they attribute to those in the assessment context may differ.

At this juncture between ethics and validity, we turn to the work of Foucault, who is among the most influential scholars to look at the intersection of truth (validity), power, and ethics. He has also resisted being categorized into one paradigm or another. Although some have mistakenly labeled him a *postmodernist* (for his own argument against this label, see Foucault, 1998, pp. 445–448), his genealogical method draws on both positivist structuralism and interpretivist hermeneutics (Dreyfus & Rabinow, 1982). Using this combination, a framework for validity and ethics can be developed that will do justice to the alternative assessment perspective, but that will also allow adherents of traditional testing and alternative assessment to communicate with each other.

To a certain extent, the issues of ethics in assessment can be thought of as aspects or expressions of power and the potential for abusing that power (i.e., through deception, the violation of privacy, or taking action concerning others without their consent). In particular, the notion of *forms of power* versus *relations of power* becomes important (Lynch & Jukuri, 1998). Foucault's (1979) earlier work, such as his history of the prison, *Discipline and Punish,* suggest an essentially negative and hopeless view of assessment as unchallengeable surveillance and control. In that work, Foucault characterizes the examination as being "at the centre of the procedures that constitute the individual as effect and object of power, as effect and object of knowledge" (p. 192). Foucault (1982) refers to the basic forms of power as *domination, exploitation,* and *subjection.* Substituting an alternative assessment procedure such as writing samples organized into a portfolio for a more traditional test or examination format does not necessarily change the power relations through which these forms are realized (McNamara, personal communication, April 11, 1998), especially if students do not control the various dimensions of the portfolio (how many products, which products, what format, how much self-reflection, etc.). In Foucault's later writing (1982, 1990, 1997a–1997e), his concept of power relations seems less dark, and it recognizes relations that resist the forms of power (Lynch & Jukuri, 1998). Foucault's later conceptualization of power relations provides a potential response to unethical assessment procedures.

Foucault's (1997) later works focus specifically on ethics, which he defines as "the relationship you have to yourself when you act," rather

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

than as "a code that would tell us how to act" (p. 131). For Foucault, ethics becomes an "aesthetics of existence" (1997d, p. 255). He accepts the Sartrean, existentialist view that the self is not given to each person but argues that the "through the moral notion of authenticity, he [Sartre] turns back to the idea that we have to be ourselves—to be truly our true self" (1997d, p. 262). This concept aligns with the portfolio as a process, where selecting and developing products and writing commentaries on their significance represent tangible access to the self-creation of a language educator (and a graduate student, and to some extent, the person as a whole).

Foucault goes on to link ethics to freedom: "For what is ethics, if not the practice of freedom, the conscious [*réfléchie*] practice of freedom? . . . Freedom is the ontological condition of ethics" (1997e, p. 284). Clearly, freedom has and always will be constrained by power. However, Foucault asserts that the forms of power are not the same as what he refers to as "relations of power" (1997e, pp. 283, 291). Domination, exploitation, and subjection are forms that result from "blocking a field of power relations, immobilizing them and preventing any reversibility of movement" (1997e, p. 283). Freeing up these immobile power relations requires actions, even political movements of liberation, to create new power relations that are reversible and mobile. These new power relations "must be controlled by practices of freedom" (1997e, p. 284); in fact, power relations cannot exist unless the participants are free to act. So, it is not a question of trying to escape from the evil of power but of creating power relations capable of maximum mobility. In the portfolio assessment context, the freedom lies in the student's ability to control and shape the portfolio process and to determine the portfolio's final form.

Foucault's sense of ethics, then, derives from considering how individuals can define themselves as a work of art, to "take care of the self" (1997c, p. 226) with an eye toward creativity rather than discovering some underlying true self. Foucault (1997e) argues that "in each [context], one plays, one establishes a different type of relationship to oneself" (p. 290). Hence, this creative activity involves the potential for creating different selves in different contexts. Again, the portfolio represents a locus of self-creation: The student presented in a portfolio will be different from the person who entered the program or the person who has a job 2 or 3 years after leaving the program.

Foucault (1997d) also establishes a link between this sense of ethics and *self writing*, which includes the ancient Greek practices known as *hupomnemata* or books of life. The description of this practice (roughly, jotting thoughts in a journal one might carry around for that purpose) has some important connections with the practice of portfolios:

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

> The point is not to pursue the indescribable, not to reveal the hidden, not to say the unsaid, but, on the contrary, to collect the already-said, to reassemble that which one could hear or read, and this to an end which is nothing less than the constitution of oneself. (p. 273)

We return to these ideas concerning alternative assessment later and synthesize them into a framework for addressing ethics and validity in portfolio assessment. In the next section, we examine the MIIS portfolio context.

THE CONTEXT

MIIS is a small (750 students), private graduate school, offering master's degrees in fields with significant cross-cultural and multilingual components (international business, international policy studies, translation and interpretation, and language education). The Graduate School of Language and Educational Linguistics offers two master's degrees, one in TESOL (MA TESOL) and one in teaching foreign languages (MA TFL).

The Program Portfolio Concept

After students have completed the required number of units in the MA TESOL or MA TFL program, they must produce a program portfolio. Students select and organize this collection of their work in accordance with program guidelines. The work selected represents the author's learning during the program, and its final form is a meaningful testament to his or her academic achievement and professional preparation. As with all portfolios, the process is important; the process allows the author to rework particular products and to reflect extensively on the learning involved. In addition, the portfolio is both comprehensive and integrative, permitting the author to explore relationships among theory, research, and practice, as well as across different aspects of educational linguistics.

The portfolio also encourages students to develop a strong professional foundation. Sections of the portfolio provide for practical products that will help students create a professional portfolio. In short, the portfolio is intended to be useful (with real outcomes), valuable both academically (the student's best work) and professionally (a strong resource for the future), and reflective.

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

History of the Program Portfolio

In February 1993, a group of students in the MA TESOL and MA TFL programs presented to the faculty a proposal for a new exit mechanism: a student portfolio, reflecting the accumulated work and learning achieved during the course of study. The portfolio would replace the existing system: a written comprehensive examination. The proposal arose principally from two sources: the students' dissatisfaction with the existing mechanism and their positive experiences with portfolios in their courses.

The ensuing negotiation culminated in a document detailing the form of the program portfolio, the process of its compilation, submission and assessment, and the assessment criteria. This document was formally adopted in August 1993 and presented to entering students as part of their orientation. The majority of this first group graduated in December 1994 and since then more than 450 students have successfully completed a program portfolio.

Goals

In the course of developing the concept and then a detailed specification of the portfolio, the faculty and students of the program articulated 11 goals:

(a) To provide clear evidence that the candidate has achieved the level of mastery implied by an MIIS degree.

(b) To integrate material and ideas from different courses in the program and to explore the relationships among theory, research, and practice.

(c) To instill a positive and productive viewpoint, resulting in greater personal investment in the learning process and a greater willingness to take risks.

(d) To stimulate clear, logical thinking and encourage habits of inquiry, taking students beyond the requirements of any particular course or assignment and creating a cumulative and integrative experience.

(e) To develop critical, creative, and independent thinking: The Portfolio format is intended to cause students to challenge orthodoxy, develop alternative explanations, break down barriers, and cross boundaries.

(f) To build a critical awareness of values; to recognize what is important in personal learning and to relate this learning to wider educational issues; to examine the intersection of intellectual, affective, and moral concerns and to pursue the implications.

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

(g) To cultivate effective writing and speaking skills: The Portfolio encourages a variety of written genres and attention to the writing process, including working with feedback, revision, editing, etc.; the Portfolio Seminar provides a forum for the presentation and discussion of ideas and for oral interaction based on responses to the writing of others.

(h) To encourage students to take responsibility for their own learning, resulting in a greater awareness of their own learning processes.

(i) To develop the skill of self-reflection and self-assessment, in terms of both the processes and products of learning.

(j) To provide the basis for a professional teaching portfolio, with the result that students develop an informed and critical focus on the range of professional opportunities and can make responsible career choices.

(k) To develop effective interpersonal skills by providing and accepting feedback, thus entering into positive interdependence with others throughout the Portfolio process. (MIIS Graduate School of Language and Educational Statistics, 2002)

Contents

Although students are not compelled to use a specific format, the portfolio contents are presented in sections (see Appendix A): Personal Statements (A), Academic Products (B), Personal, Practical, and Professional Products (C), and Critiques (D), including self-critiques. This system separates the academic material (largely in Section B) from the professional elements (Section C), following an introduction (Section A, the Commentary and Position Paper) that sets out the portfolio's key themes and provides a context for what follows in terms of the author's own goals, experiences, and ideas. The portfolio normally concludes with the assessment section (D), which contains contributions from the author, the author's peers, and the faculty reviewers.

The Process

The following process has been established for the program portfolio:

1. At their orientation, new students are introduced to the concept of the Portfolio and given a copy of the document. Three or four weeks later, they will be invited to a meeting to discuss the details and to inspect Portfolio samples.
2. In approximately the twelfth week of their first semester, students attend another meeting where they develop a plan for working on the Portfolio over the semester break and in the early part of the second semester.

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

3. In subsequent semesters, as students add to their Portfolios from current projects, assignments, and activities, another meeting is held and students discuss with one another the current contents of their Portfolios and plans for revising items.
4. In their final semester, students register for a two-unit course Portfolio Seminar. Under faculty guidance, Portfolio items are presented and accorded critical feedback, leading to the final version of the Portfolio.
5. Each Portfolio is then examined by two faculty members, who provide feedback in writing. Before issuing their final clearance, these readers hold a conference with the candidate to discuss selected aspects of the work in the Portfolio. In particular, the readers' reservations concerning particular Portfolio items are addressed either through discussion or by revision of written products (or both), as reflected in the following judgment categories:

    Pass (No written revisions required; optional conference questions);

    Pass, pending successful Portfolio conference;

    Pass, pending successful written revision;

    Pass, pending successful Portfolio conference and written revision;

    Not pass (Candidate must revise Portfolio for the next submission deadline).

6. Once cleared for graduation, students display their finished product at a program-wide event where peers may share one another's successes; students at earlier stages in the Program can browse to gain insights into the Portfolio process. (MIIS Graduate School of Language and Educational Linguistics, 2002)

Particularly noteworthy is that the process is initiated within hours of the students' arrival at MIIS, and its importance is stressed at regular intervals. Students are therefore aware of their responsibilities at the outset of the program, and their awareness of what they can or should be doing is regularly maintained through informational meetings. The portfolio seminar is also a key feature, providing information, support, facilitation, and feedback during the weeks when students are assembling their portfolios. The need for such a process has become formalized in, for example, requiring students to submit two drafts of the position paper along with the final version and strong evidence, in Section D2, of peer feedback both given and received.

Assessment Criteria

The program portfolio is subject to continuous assessment, beginning with the student's earliest ideas during the first semester and culminating

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

with the final form, submitted toward the end of the final semester. The assessment comprises three facets: self-assessment by the author, peer assessment by fellow students, and faculty assessment by the two readers. From the process of student-faculty consultation, eight criteria were developed to guide the assessment; they were subsequently revised in the light of reading and responding to the first four sets of portfolios. They are

Criterion 1: Depth (including critical synthesis and adequate coverage of the literature)

Criterion 2: Accuracy

Criterion 3: Contribution to Personal and Professional Development

Criterion 4: Voice

Criterion 5: Breadth

Criterion 6: Integration (of theory, research and practice)

Criterion 7: Contribution to the Profession (actual or potential)

Criterion 8: Writing and Presentation (MIIS Graduate School of Language and Educational Linguistics, 2002)

These criteria (see Appendix B for descriptions) are applied, where appropriate and relevant, at each stage of assessment.

THE FRAMEWORK: ANALYZING VALIDITY AND ETHICS OF PORTFOLIO ASSESSMENT

Considering validity from the alternative assessment perspective and Foucault's thought in relation to ethics, then, has provided a framework for examining portfolio assessment in the MIIS programs context.

Fairness

Most validity frameworks, including those from the positivist perspective, include some notion of *fairness*. However, fairness in traditional perspectives can differ significantly from fairness in alternative perspectives. In the traditional formulation, with its focus on measurement, fairness is chiefly concerned with making sure everyone is treated the same way—that everyone receives the same assessment procedure, and that the results of that assessment are interpreted in the same way for all candidates. Another way of asking the fairness of interpretation ques-

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

tion: Do the assessment results depend on who is judging or on which task or portfolio component is being assessed? From the traditional, measurement-based perspective, assessment results should be independent of rater and task. From the alternative perspective, assessment results reflect the different knowledge and understanding that different raters bring to the assessment process and the different tasks that the students complete, each of which reveals a different aspect of the students' knowledge and abilities. Whereas traditional measurement sees interrater agreement as a sign of fairness (or as it is traditionally known, reliability), alternative assessment accepts disagreement as a valid part of the process. That is, raters may disagree for the right reasons or agree for the wrong reasons.

Writing about portfolio assessment in college composition programs, Elbow (1991) makes this point:

> When all writing is alike, it is easier to agree about it. What a mess portfolios make, then, for psychometricians looking for reliable scoring, what a problem for holistic testers who've been bragging to professional psychometricians about good reliability scores on holistic readings. But in this very difficulty about reliability, we see another benefit of portfolios—another reason why portfolio assessment appeals to many thoughtful people in our profession. For the truth is that many of us cannot help feeling that if reliability is high—if readers all agree about the worth of a piece of writing—something must be fishy. (p. iv)

In this view, therefore, it is preferable for readers to behave as themselves, thus mirroring the writer's sense of ontological authenticity. Indeed, Elbow describes this claim for portfolios as *mirror validity*; if validity usually involves judging how well the evidence represents the student's actual learning and progress, then mirror validity involves judging how well the assessors look at that evidence: Does the assessment practice appropriately accommodate how readers actually read?

Although both perspectives may seek ultimate consensus on the meaning of a student's portfolio, they will likely achieve that consensus using different processes. From the alternative assessment perspective, reaching a fair consensus on the meaning of a student's portfolio (as opposed to a statistically significant correlation between the judgments of two or more raters), involves a consideration of the following questions:

- Does the portfolio assessment process take into account the perspectives of all affected participants? Is each portfolio a true reflection of its author's learning? And is the assessment outcome (rating, feedback, subsequent dialogue) a true reflection of the reader's natural reactions?

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

- Is the assessment structured so that the relations of power are mobile, reversible, and reciprocal, thus maximizing ethical behavior in Foucault's sense?

Ontological Authenticity

This category of ethics and validity comes from Guba and Lincoln's (1989) *authenticity criteria*. For Guba and Lincoln, ontological authenticity means being able to access and use information from the research (or assessment) in a meaningful way. This idea intersects with Foucault's "care of self" (1997c, p. 226) and "practices of the self" (1997c, p. 225), practices that actively constitute an identity or, more accurately, identities for the self. Guba and Lincoln's notion can be extended to include this ability to establish a meaningful identity. *Ontologically authenticity*, then, is a characteristic of ethical and valid assessment that asks, do the participants in the assessment process establish a meaningful identity, a sense of who they are? This category, like the others, includes all participants and stakeholders in the assessment process: teachers, assessors, administrators, parents, and community members. Using Foucault's thought emphasizes creativity, the active construction of identity, rather than authenticity per se (*authenticity* in the Sartrean sense, as discussed in Foucault, 1997d).

Cross-Referential Authenticity

This category is derived from Guba and Lincoln's (1989) *educative authenticity*, which examines the degree of understanding gained about other's perspectives. *Cross-referential authenticity* examines the understanding of the identities that others have constituted for themselves as a result of the assessment process. This category asks, are the participants in the assessment process able to improve their understanding of perspectives from outside their own group? For example, do students understand teachers better? Does the teacher from a different class understand a colleague and/or her students better?

Impact or Consequential Validity

This type of validity derives from Bachman and Palmer's (1996) sense of *impact* in their test-usefulness framework. It corresponds also to Messick's (1989) *consequential validity* (the value implications and social consequences of test interpretation and use), and to Guba and Lincoln's

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

(1989) *catalytic authenticity* (the degree to which something happens as a result of the research or assessment process). In our framework, accordingly, ethical and valid assessment examines the outcomes, intended and unintended, of the assessment process. The question to be examined is, what is actually done as a result of the assessment? For example, is the curriculum changed in some way, either its content or instructional practices or its assessment practices at the individual course level? Do teachers-in-development modify some aspect of their approach to teaching? How do graduates of the program present themselves to prospective employers?

Thinking about this question a bit further reveals the link between values and consequences (highlighted by Messick, 1989). After the assessor has identified that something has been done as a result of the assessment, he or she still must determine whether this consequence, this impact, is good or bad. To remain consistent with the other categories of this framework, the evaluation of impact should include the multiple perspectives that make up the assessment setting. It would be a negotiated consensus of some sort.

Evolved Power Relations

To a certain extent, this category overlaps with Guba and Lincoln's *tactical authenticity* (1989) or the degree to which participants are empowered to carry out the changes that the research or assessment process makes possible. Adding Foucault's notion of power relations, which are central to determining ethics, creates further overlap with the other categories in this framework. For example, determining ontological authenticity would clearly involve examining power relations, and one important, potential impact or consequence of the assessment process would be a change in power relations. However, because Foucault argues that power is an important element in free and ethical relations, I use it in our framework as a category of particular focus. This category asks the following questions:

- Does the assessment process cause participants to change how they relate to each other and to themselves? For example, do the students assume and obtain more responsibility for the curriculum? Do the teachers gain control over assessment policies previously established by others?
- Do these changes become fixed, or are they established as reversible, mobile relations of power?

In the same way that impact must be evaluated, a change in the way participants relate to each other also must be evaluated as positive or

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

negative. The answer to the second question, however, includes its own basis for evaluation: Power relations that are reversible and mobile are valued as good.

Table 1 contains a summary of the framework for considering ethics and validity from the alternative assessment perspective. It also shows the framework's relationship to the MIIS portfolio.

THE MIIS PORTFOLIO AND THE ALTERNATIVE ASSESSMENT FRAMEWORK

This section considers the MIIS portfolio in light of the proposed framework's five categories. The investigation used three data sets. The first data set, and the most substantial, was copies of the Commentary and Self-Assessment sections from the portfolios. These two sections were most likely to contain students' comments on the portfolio as a whole and its process. As part of a long-term (1997 to the present) investigation of validity and other assessment issues, each portfolio author was asked to provide a copy of these documents and to sign a release form permitting their inclusion in the study. The findings presented in this article were drawn from a corpus of 172 commentaries and 168 self-assessments. The second data set was recordings of face-to-face interviews with 23 program graduates and notes from telephone interviews with 29 graduates. All these interviews took place within 36 months after graduation. Participants were asked to reflect back on the portfolio and its process, describe perceived strengths and weaknesses, and to describe how they have used the portfolio in their professional careers. The third data set was audio recordings of portfolio conferences, made mostly during the first 4 years (1994–1998) of portfolio activity to better understand this aspect of the process. In the present study, we became especially interested in students' comments during the final few minutes of a conference, when students might be asked about their experience with the portfolio as a whole.

These data sets were examined for comments on ethics and validity issues. The extracts containing these comments were then coded according to which of the five categories in the ethical and validity framework they most directly addressed. This process produced approximately 600 passages, of which 540 (90%) were positive (i.e., comments suggesting that the category's ethical and validity concerns had been met), about 40 were neutral (i.e., comment relating to the category, but not clearly positive or negative), and 20 were negative (i.e., comments suggesting that the category's ethical and validity concerns had not been met). Space restrictions limit the number of passages that we can discuss in this

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

TABLE 1

The MIIS Program Portfolio: Ethics and Validity

Ethics-validity category	Definition	Extension to alternative assessment and portfolios	Relevant portfolio sections[1]	Examples of data[2]	Assessment criteria[3]	Goals[4]
Fairness	Everyone receives the same assessment procedure, and the results of that assessment are interpreted in the same way for all candidates; all participants in a particular assessment setting have been taken into account.	Are the perspectives of all affected participants in the portfolio assessment process being taken into account?	A1, A2, D1	• Role of portfolio seminar • Faculty discussion • Genesis and development of portfolio system	1, 2, 5	c, f, g
Ontological authenticity	Participants gain and use information as a result of the assessment process.	Do the participants in the assessment process establish a meaningful identity, a sense of who they are?	A1, C1, D1, cover notes	• Portfolio author data • Portfolio seminar	3, 4, 7, 8	c, e, h, i
Cross-referential authenticity	Stakeholders and participants gain an understanding of the perspectives and meaning constructions of those outside their own group.	Are the participants in the assessment process able to gain an improved understanding of the perspectives outside their own group?	B1, B2, B3	• Faculty review process and portfolio conference • Faculty have better understanding of program as a whole	5, 6	d, f, j, k

(continued on p. 282)

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

TABLE 1 *(continued)*

The MIIS Program Portfolio: Ethics and Validity

Ethics-validity category	Definition	Extension to alternative assessment and portfolios	Relevant portfolio sections[1]	Examples of data[2]	Assessment criteria[3]	Goals[4]
Impact	A consideration of the values and goals that inform assessment interpretations, and the consequences of interpreting and using assessment information; the degree to which something happens as a result of the assessment process.	What is actually done as a result of the assessment?	B1, B3, C2, C3, C4	• Curriculum integration begins • Graduates use portfolios in job interviews • Changes in feedback to student work	3, 5, 6, 7	a–k
Evolved power relations	The degree to which participants are empowered to carry out the changes that are made possible through the assessment process.	Do the participants change the way they relate to each other and to themselves?	A1, C4, D1, D2	• Genesis and development of portfolio system • Portfolio conference	3, 4, 6, 7, 8	c, e, h

[1] See Appendix A.
[2] For student quotations, see relevant sections.
[3] Criteria numbers: 1. Depth, 2. Accuracy, 3. Contribution to Personal and Professional Development, 4. Voice, 5. Breadth, 6. Integration, 7. Contribution to the Profession, 8. Writing and Presentation. See Appendix B for definitions.
[4] For definitions of goals, see pp. 273–274.

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

article. In the interests of balance, neutral and negative comments are overrepresented; otherwise, the comments used to illustrate the five categories are representative of the corpus. Although not all students reported a wholly positive experience with the portfolio, we did find a bountiful supply of remarks supporting and detailing the categories.

In the next section, we obtain a preliminary indication of the match with the ethical-validity framework by examining each ethical-validity category in terms of the structure, goals, and criteria that define the portfolio, as well as aspects of the process of compiling and completing it. We then offer evidence from the three data sets mentioned earlier. This approach to validity examines the MIIS portfolio process and the students' portfolio products to see whether they support the categories in the framework. This approach also builds a validity argument (cf. Chapelle, 1998), or an argument, from the alternative assessment perspective, for the credibility or trustworthiness of the portfolio assessment process and the decisions or judgments made from it.

Fairness

The MIIS portfolio's structure and process provide multiple sources of evidence for assessment. This multiplicity should facilitate the desired inclusion of all participants' perspectives, especially the students and the portfolio readers. Thus, the portfolio has multiple sections and various subsections (see Appendix A). The portfolios have multiple goals (11) and multiple assessment criteria (8). These multiple opportunities to display learning and the multiple goals and criteria being used to judge that learning provide the fairness defined by this category. Similarly, the portfolio seminar process facilitates multiple drafts of key elements, especially of the position paper, two earlier drafts of which must be included in the portfolio, along with feedback from the portfolio readers. The consolidation of all these dimensions in the portfolio document means that all authors are bound by the same system, yet there is considerable room for individual variation within that system, and the perspectives of students and portfolio readers will be included.

Aspects of the way that the MIIS portfolio came into being indicate its potential for fairness. The portfolio system at MIIS was based on a student initiative, followed by a detailed student proposal and subsequent faculty-student discussion. This process suggests that the resulting system has built into its core the potential for the reciprocal power relations characterized earlier as a part of ethical assessment and fairness. The portfolio seminar, which allows students to raise issues, possibilities, and questions and where the instructor acts as a liaison between students and faculty, forwarding questions and issues to colleagues

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

and taking the consensus response back to the students, suggests that multiple participant perspectives are being taken into account and that power relations are reciprocal.

A closer look at the goals of the MIIS portfolio also suggests the potential for fairness. Goal (e), with its aim of "independent thinking" and encouraging students to "question authority, break down barriers and cross boundaries," should foster reciprocal power relations. Similarly, Goals (f) and (g) encourage students to take responsibility for their own learning and engage in self-reflection. Specific sections of the MIIS portfolio also suggest an underlying concern with fairness: Section A1, Commentary, where authors express key personal ideas and issues, including their theme; and Sections D1, Self-Assessment, and D2, Peer Critiques, where student authors and their peers take responsibility for part of the portfolio assessment.

Does the evidence show that the MIIS portfolio is realizing its potential for fairness? Students comments relating to this category appeared in the student portfolio writings, especially in the Commentary and Self-Assessment sections. A number of authors, for instance, expressed satisfaction with the opportunity to represent their own perspective and construct their own meaning. Here is one example from a self-assessment:

> I think this method of assessment turned out to be a perfect way to account for my own growth and assess in what ways and how well I meet my own goals. Overall, I feel that I am very close to dealing with all issues I wanted my portfolio to represent or address: due to this accomplishment, I am completely satisfied with how my wants manifested into the form of my Portfolio.

Another student made a similar point in a self-assessment, comparing the portfolio with more traditional forms of assessment:

> The main strength of the process was the actual enjoyment I got out of trying to clarify and synthesize all the new ideas I'm learning. I have never had the opportunity to go back at the end of the semester and reflect on what I had learned. I usually walked out of the final exam and promptly forgot everything. This process is much more relevant in producing a final product that is a true reflection of learning than the process of studying for the product of a final exam.

In contrast, some students felt that the portfolio did not truly reflect their learning across all sections, often because of time constraints, as the following excerpt from a self-assessment shows:

> Among all the sections of the portfolio, I am most satisfied with the position paper. I enjoyed the process of writing it and I am happy with the final product. I am least satisfied with my B1 Revised Projects, which I must admit

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

suffered somewhat from the inordinate amount of time I put into the position paper.

Some students occasionally struck a more negative note, as in this extract from an exit interview:

> One of my main concerns . . . was that the professors sometimes seemed to pay more attention to whether or not my reference lists were in APA format than they were to the content and my ideas. The bottom line is that some of the creative essence of the portfolio was lost in the process. There needs to be less emphasis on small insignificant details and more emphasis on true learning and creative expression.

Ontological Authenticity

The portfolio and its process address this category of validity in those aspects that enable participants to establish a meaningful identity. These features of the portfolio and its process reflect the potential for ontological authenticity:

- Commentary section, where authors often preview their new or changing identity
- Personal Writing section, where this identity is illustrated in an especially personal voice
- Self-Assessment section, where the identity is reviewed and its strengths and weaknesses assessed
- cover notes, which often tie a particular portfolio item to one or more aspects of the author's identity
- portfolio seminar, where activities early in the seminar—brainstorming, early drafts of key pieces, exploring ideas and themes, and early, informal, and exploratory expressions of identity—foster self-awareness and identity development. The seminar instructor also reminds participants that all portfolios are unique, that is, ontologically authentic.
- The following portfolio goals promote a profound sense of self in the portfolio authors:
    - (c) personal investment
    - (e) independent thinking
    - (h) responsibility and awareness
    - (i) self-reflection and self-awareness
- The following assessment criteria create the potential for ontological authenticity:

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

Criterion 3, contribution to personal and professional development, requires authors to specifically exhibit facets of growth.

Criterion 4, voice, stresses the need for readers to hear the individual behind the work.

Criterion 7, contribution to the profession, includes the author's identity as a teacher, along with the potential for professional contributions beyond the classroom.

Criterion 8, writing and presentation, focuses on the academic self and the qualities required for a successful graduate student.

A person entering the program tends first to become a *graduate student* and subsequently explores what kind of a student she is or wants to be. Similarly, most decide to be a *language teacher* and then later decide what kind of a teacher they might be. The portfolio process facilitates that awareness and then formalizes the outcome, as this quotation from a postgraduation interview shows:

> Q: Do you think the portfolio is a relevant, good way to present yourself as a professional?
>
> A: Of course I do! The portfolio represents who I am, what I do, and what I believe. Because it is basically a never ending project, it shows the growth and changes I have endured, as I have had more and more experiences. I am constantly growing as a Teacher, as well as a person, therefore the portfolio continues to grow and change. I think a portfolio is a great way to show our peers how much we have grown.

This extract from a self-assessment shows the multifaceted nature of the author's identity:

> Whether there be places I might have elaborated or improved, I feel the portfolio is a process, and, by nature, is never totally complete. I am certain, though, that this portfolio, as it exists today as you read it, succeeds in at least one way. This portfolio represents me in all my manifestations: M. the linguist, M. the theorist, M. the writer, M. the student, and, most importantly, M. the teacher.

The data also contained a handful of reservations, such as the following from a self-assessment, that focused on the difficulty of finding room for more personal or practical documents:

> I would have liked to include more journal writing and reflections of my teaching as well as classroom observation reports I have written, to illustrate features of my approach to teaching. However, the academic demands of the portfolio made this impossible.

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

Cross-Referential Authenticity

This feature, the improved understanding of others' perspectives, can be approached from either the students' or the teachers' point of view. The following observations deal with the students' viewpoint:

- The crucial part of the portfolio is Section B, and especially B1, where the two revised projects incorporate input from several different classes, and, thereby, different instructors. These revisions thus involve reviewing and reconciling the goals and approaches of different instructors as well as the context of those classes.
- Assessment Criterion 5 (breadth) and Criterion 6 (integration) require students not to ignore any key components of the field and to synthesize concepts, data, and procedures into a meaningful whole.
- Several portfolio goals are relevant to cross-referential authenticity:

  Goal (d) stresses "a cumulative and integrative experience."

  Goal (f) posits the significant "intersection of intellectual, affective and moral concerns."

  Goal (j) underscores the requirement that students represent their understanding of teaching and present their own "informed and critical focus" on the profession.

  Goal (k) represents cross-reference to the peer group and the importance of "interpersonal skills" and "positive interdependence with others."

The following quotation from a portfolio commentary shows how the portfolio process enabled a student to understand more about teachers and teaching:

> My own belief is that a teacher is also a learner, and, therefore, thinking about what learning involves is essential to being a teacher. I see learning as a cycle of interaction, introspection and innovation. . . . I think it is enough to say here that every project contained in this portfolio required me to work through the cycle numerous times and represents my own learning process. While I have primarily played the role of learner in this program, I have approached every project with the mind of a teacher. Therefore, this portfolio represents me as both. I cannot imagine it any other way.

The final sentence of this abstract emphasizes how powerful this aspect of validity can be for some students.

From the faculty perspective, three points stand out:

- The assessment process: Two faculty readers examine and respond to the portfolio. When, for example, an SLA specialist and an assessment

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

guru read the same portfolio, they must interact to agree on an assessment outcome, a process that most faculty participants regard as fruitful and valuable. Similarly, the conversation in the portfolio conference among the portfolio author and the two readers can be professionally enriching.

- The process has helped faculty readers to better understand the program as a whole and the contributions of individual components and colleagues. Reading papers and projects from courses one does not teach provides valuable insights into the program.
- The process enables faculty to understand more fully the experience of being a student in the program. Sections A1 and D1 are particularly valuable because students discuss the challenges that they have met and overcome in their work and learning. Faculty have manifested their new understanding of students' experience in, for example, how they respond to student work.

Although most comments in the data were positive, some indicated a shortfall in achieving full understanding of other perspectives, especially the professional:

> Blending the theoretical and practical has given us the opportunity to synthesize inter-weaving relationships [in education] on personal, professional and global levels, and to form lucid and tangible ideas of how language and culture education relate to my life and my goals as a teacher. However, this cross-referencing is curtailed by the emphasis in the portfolio on the accurate representation of theory. My teaching goals have been somewhat lost in the process. (From a self-assessment)

Impact

This feature, highlighting what is actually done as a result of the assessment, can be illustrated in two forms: effects within the program and effects beyond the program. Summing these two sets of consequences reveals that impact is perhaps the most powerful of the five categories of validity explored in this article. For example, all 11 portfolio goals are related to impact, whether on the author as a person (e, f, h, i, k), an academic (a, b, d, e, g, i), or a professional (b, j, i, k), or, more broadly, on the program itself (b, d, e) and on the profession (g, j, k). In the case of effects within the program, the following trends are noteworthy:

- The portfolio has led to various attempts to integrate the MA TESOL curriculum. Part of the impetus for this change came from students, who posed this question: Because we are required to integrate material in our portfolios (see Criterion 6), why isn't the program

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

itself more thoroughly integrated? The response has included creating joint projects across two or more classes and a series of Integration Events where, for example, faculty and student panels respond to issues involving two or more specialties (and, therefore, two or more courses). The portfolio process compels faculty readers to pay close attention to breadth (Criterion 5) and integration (Criterion 6) and thence transport new insights back to their syllabi and teaching practices.

- Faculty responses to course projects have changed: Rather than treating a course product as a fixed, terminal entity, faculty feedback now includes suggestions for reworking (expanding or combining with other work), should the piece be selected for inclusion in Section B1, Revisions.
- Some classes require students to summarize products, which helps them to develop this important skill early, well before they have to create portfolio Section B3, Summaries of Papers and Projects.

From the student perspective, the key point is undoubtedly the portfolio's professional development aspect. Criterion 3, contribution to personal and professional development, and Criterion 7, contribution to the profession, speak directly to this aspect, and Goals (b), (d), (i), (j), and (k) are all important correlates. In terms of portfolio contents, Sections C2, C3, and C4 are the most important because they will form the core of a professional portfolio. Some students, in fact, report that they have designed their portfolios, in some cases down to such details as font choices, with the professional product in mind. One student included a specific section titled "Transition to a Professional Portfolio." Another made Sections C2, C3, and C4 function as a portfolio within a portfolio, a professional embryo detachable from the whole.

The students' attitude toward this professional component is summed up in this extract from a self-assessment:

> I consider this portfolio not the solved puzzle of my philosophy as it will always apply to my teaching, but rather I see it as the starting point for my career: it is the beginning, not the end.

The data contain a scattering of comments, however, that are less enthusiastic, as this extract from a self-assessment shows:

> Because of the detailed requirements, the multiple drafts and the time pressure, it was all too easy to lose sight of one's ultimate goals while compiling the portfolio. Throughout the process, I tried to remind myself that I came here to be a better teacher. However, at this point, it is hard to see the impact of the portfolio on my future teaching.

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

Evolved Power Relations

The portfolio process reflects this category, the empowerment of participants to carry out the changes made possible by the assessment procedure, in how it changes the way the various parties relate to each other and to themselves. For example, portfolio author statements often manifest Goal (h), taking responsibility for one's own learning, by quoting others. Other elements that suggest evolving power relations include the following:

- the portfolio conference, where a newly hatched language education professional meets with two new colleagues to discuss issues of common interest
- Sections D1 and D2, which show evidence of the multiple partnerships in the assessment process
- Criterion 3, personal and professional development
- Criterion 4, strong personal voice
- Criterion 7, professional contributions

The following extract from a portfolio commentary shows that the way students relate to themselves and others changes during the portfolio process:

> I believe that all learners have the right to ask, why do I learn what I learn? After many years of being a learner in many different classroom settings, I now realize that I, like all learners, have this basic right. I don't always know how to respond to this question, but the search for answers empowers me and enlightens me about the learning process and my own learning behavior. This portfolio documents some of the answers I have found. It reveals what I feel and believe about myself and about learning in general. Now I know that I am able to reflect and (in John Dewey's words), "summarize . . . a developing experience." I now believe in myself as a resourceful and creative professional, and I am more aware of my skills as a critical thinker. I know I am a willing and active participant in the learning process. I am convinced that my learning is more valuable and meaningful because I have exercised the right to ask, why do I learn what I learn? This portfolio is proof that I am able to take responsibility for my learning and my development as a thinking individual.

Although all the authors incorporated the stated goals and assessment criteria, a number of them also made it clear that they expected their work to be assessed on its own terms, as in this final paragraph from a portfolio commentary:

> Power, for this portfolio, is defined as the ability to organize and maintain a state of "flow consciousness" (Csikszentmihalyi, 1992) under conditions of

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

high challenge, met by high levels of developing skills, in the pursuit of personally meaningful goals. . . . This portfolio is concerned with issues relating to 1) one individual's efforts to organize her own state of consciousness to meet flow conditions in the domain of learning in the MIIS setting and 2) this individual's desire to help others who share the desire to learn in a flow process. Underlying assumptions are: that organized states are preferable to disorganized states and are attainable; that "flow" states are preferable to "stuck" states; that flow may be experienced by some individuals with a certain skill level through a process involving awareness and the right efforts; and that flow conditions are synonymous with optimal learning conditions. Not everyone will agree with these assumptions, but they are the basis of all the work and play that follow. So this portfolio is for those who share this text's assumptions. For those who don't, let them create their own texts!

In general, students who had been through the process continued to offer suggestions for improving the portfolio, as, for example, in this exit interview:

> With regard to the size, the scope and amount of time involved in the portfolio process is overwhelming. One suggestion is to require one B1 revision instead of two. This would cut down on the amount of time and stress and let students focus on one project, enjoy it more. . . . The excessive requirements and details stifled my creativity.

The faculty are currently considering precisely this reduction in the scope of Section B1.

The origin of the program portfolio is also evidence that it fosters evolved power relations. A quotation from the February 1993 student proposal to the faculty illustrates this point:

> The process of creating a portfolio instills a positive and productive mindset. We believe that the adoption of the following proposal, or a revision thereof, would positively benefit student learning by accomplishing the following: stimulating and encouraging habits of enquiry, fostering the skill of critical and creative thinking, promoting an awareness of values, and cultivating both effective writing and speaking skills. Moreover, students will share a greater responsibility for learning, developing skills of independent thinking and self-assessment, resulting in a higher degree of personal investment and willingness to take risks. They will have a better awareness of their learning processes, the reasoning behind their own opinions, and possibly what path they would like to pursue after graduation.

CONCLUSION

Portfolios seem to offer an excellent means of creating ethical and valid assessment practices, analogous to the Greek books of life. Through

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

their work on portfolios, students address ontological authenticity; that is, they answer the questions of academic and professional identity: Am I a successful member of the MIIS program? Am I really an ESL teacher? Am I a member of the TESOL community? The portfolio, like the *hupomnemata,* involves reflection and an active piecing together of ideas and perceptions, a "formation of the self out of the collected discourse of others" (Foucault, 1997b, p. 217). This active constitution of self is similar to the actions of the protagonist in Camus's *The Stranger.* In prison, unlike his previous existence, he begins to reflect on things, even polishing his metal meal plate so that he can see himself (prior to the murder and jail sentence, he had no mirror in his apartment). This polishing is the metaphorical work of assembling a portfolio; once it is completed, you have something in which you can see yourself, a reflection on which you can reflect further. You have tackled the ontological work of establishing your identity.

In the setting that we have examined in this article, students used the portfolio to create their professional identities and to consolidate their academic identities. In many cases, students expressed these identities in ways that demonstrated an integration with their personal identity as well. This fact created a sensitive situation in the assessment process because criticizing portfolio elements was likely to make students feel that they were being personally attacked. The assessment process can be seen as ethical and valid, however, to the extent that the students had reciprocal rights to make observations and judgments—a key aspect of fairness, as we have used the term. It is also important to determine whether the teachers or the assessors came to a fuller understanding of the student's construction of identities, and whether the student came to a fuller understanding of the teacher's or assessor's constructed identities—what we have called cross-referential authenticity. That is, did both of these groups of participants come to a better understanding of what it means to be a teacher in this field?

However, teachers and assessors do not have a completely passive role, either in the teaching or the assessing process. As Foucault (1997e) expressed it,

> I see nothing wrong in the practice of a person who, knowing more than others in a specific game of truth, tells those others what to do, teaches them, and transmits knowledge and techniques to them. . . . [The problem] is knowing how to avoid the kind of domination effects where a kid is subjected to the arbitrary and unnecessary authority of a teacher, or a student is put under the thumb of a professor who abuses his authority. (p. 299)

Examining impact and evolved power relations, in particular, will show whether the teaching and learning have proceeded ethically and validly. This thinking suggests that power relations are not inherently evil

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

and need not be escaped. They are, rather, part of the human condition, and educators must work to develop practices—in this case, practices of assessment—that maximize the mobility, the reciprocality of those relations.

THE AUTHORS

Brian Lynch is an associate professor in the Department of Applied Linguistics at Portland State University, Oregon, USA. He has published in the areas of language testing, program evaluation, languages for specific purposes, and computer-assisted language learning.

Peter Shaw is a professor in the Graduate School of Language and Educational Linguistics at the Monterrey Institute for International Studies, California, USA. His research interests include language teaching pedagogy, co-operative learning, curriculum theory, content-based instruction, and discourse analysis.

REFERENCES

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice.* Oxford, England: Oxford University Press.

Birenbaum, M. (1996). Assessment 2000: Towards a pluralistic approach to assessment. In M. Birenbaum & F. J. R. C. Dochy (Eds.), *Alternatives in assessment of achievements, learning processes, and prior knowledge* (pp. 3–29). Dordrecht, Netherlands: Kluwer Academic.

Brown, J. D., & Hudson, T. D. (1998). The alternatives in language assessment. *TESOL Quarterly, 32,* 653–675.

Chapelle, C. A. (1998). Construct definition and validity inquiry in SLA research. In L. F. Bachman & A. D. Cohen (Eds.), *Interfaces between SLA and language testing research* (pp. 32–70). Cambridge, England: Cambridge University Press.

Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review, 64,* 5–30.

Dreyfus, H. L., & Rabinow, P. (1982). *Michel Foucault: Beyond structuralism and hermeneutics.* Brighton, England: The Harvester.

Elbow, P. (1991). Preface. In P. Belanoff & M. Dickson (Eds.), *Portfolios: Process and product* (pp. i–iv). Portsmouth, NH: Heinemann/Boynton-Cook.

Foucault, M. (1979). *Discipline and punish: The birth of the prison* (A. Sheridan, Trans.). New York: Vintage.

Foucault, M. (1982). The subject and power. In H. L. Dreyfus & P. Rabinow, *Michel Foucault: beyond structuralism and hermeneutics* (pp. 208–226). Brighton, England: The Harvester Press.

Foucault, M. (1990). *The history of sexuality: Volumes 1, 2, and 3* (R. Hurley, Trans.). New York: Vintage.

Foucault, M. (1997a). Michel Foucault: an interview by Stephen Riggins. In P. Rabinow (Ed.), *The essential works of Michel Foucault, 1954–1984, Vol. 1: Ethics: Subjectivity and truth* (pp. 121–133). London: Allen Lane.

Foucault, M. (1997b). Self writing. In P. Rabinow (Ed.), *The essential works of Michel Foucault, 1954–1984, Vol. 1: Ethics: Subjectivity and truth* (pp. 207–222). London: Allen Lane.

Foucault, M. (1997c). Technologies of the self. In P. Rabinow (Ed.), *The essential*

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

works of Michel Foucault, 1954–1984, Vol. 1: Ethics: Subjectivity and truth (pp. 223–251). London: Allen Lane.

Foucault, M. (1997d). On the genealogy of ethics: An overview of a work in progress. In P. Rabinow (Ed.), *The essential works of Michel Foucault, 1954–1984, Vol. 1: Ethics: Subjectivity and truth* (pp. 253–280). London: Allen Lane.

Foucault, M. (1997e). The ethics of the concern of the self as the practice of freedom. In P. Rabinow (Ed.), *The essential works of Michel Foucault, 1954–1984, Vol. 1: Ethics: Subjectivity and truth* (pp. 281–301). London: Allen Lane.

Foucault, M. (1998). Structuralism and poststructuralism. In J. D. Faubion (Ed.), *The essential works of Michel Foucault, 1954–1984, Vol. 2: Aesthetics, method, and epistemology* (pp. 443–458). New York: The New Press.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation.* Newbury Park, CA: Sage.

Hamayan, E. V. (1995). Approaches to alternative assessment. *Annual Review of Applied Linguistics, 15,* 212–226.

Hamp-Lyons, L. (1989). Language testing and ethics. *Prospect, 5*(1), 7–15.

Hamp-Lyons, L., & Condon, W. (2000). *Assessing the portfolio: Principles for practice, theory, and research.* Cresskill, NJ: Hampton.

Hamp-Lyons, L., & Lynch, B. K. (1998). Perspectives on validity: A historical analysis of language testing conference abstracts. In A. J. Kunnan (Ed.), *Validation in language assessment* (pp. 253–276). Mahwah, NJ: Lawrence Erlbaum.

Kincheloe, J. L., & McLaren, P. (2000). Rethinking critical theory and qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The handbook of qualitative research* (2nd ed., pp. 279–313). Thousand Oaks, CA: Pine Forge.

Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15–21.

Lynch, B. K. (2001). Rethinking assessment from a critical perspective. *Language Testing, 18*(4), 351–372.

Lynch, D. A., & Jukuri, S. D. (1998). Beyond master and slave: Reconciling our fears of power in the writing classroom. *Rhetoric Review, 16*(2), 270–288.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). Washington, DC: The American Council on Education and the National Council on Measurement in Education.

MIIS Graduate School of Language and Educational Linguistics. (2002). Program level assessment in the MATESOL/MATEFL curriculum: The portfolio. [Departmental document]. Monterey, CA: Monterey Institute of International Studies.

Moss, P. A. (1994). Can there be validity without reliability? *Educational Researcher, 23*(2), 5–12.

Moss, P. A. (1996). Enlarging the dialogue in educational measurement: Voices from interpretive research traditions. *Educational Researcher, 25*(1), 20–28.

Norris, J. M., Brown, J. D., Hudson, T., & Yoshioka, J. (1998). *Designing second language performance assessments.* Honolulu, HI: University of Hawai'i Press, Second Language Teaching and Curriculum Center.

Pennycook, A. (2001). *Critical applied linguistics: A critical introduction.* Mahwah, NJ: Lawrence Erlbaum.

Phillips, D. C. (1990). Postpositivistic science: Myths and realities. In E. G. Guba (Ed.), *The paradigm dialog* (pp. 31–45). Newbury Park, CA: Sage.

Schwandt, T. D. (2000). Three epistemological stances for qualitative inquiry: Interpretivism, hermeneutics, and social constructionism. In N. K. Denzin & Y. S. Lincoln (Eds.), *The handbook of qualitative research* (2nd ed., pp. 189–213). Thousand Oak, CA: Sage.

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

Shohamy, E. (1996). Language testing: Matching assessment procedures with language knowledge. In M. Birenbaum & F. J. R. C. Dochy (Eds.), *Alternatives in assessment of achievements, learning processes, and prior knowledge* (pp. 143–160). Dordrecht, Netherlands: Kluwer Academic.
Wile, J. M., & Tierney, R. J. (1996). Tensions in assessment: The battle over portfolios, curriculum, and control. In R. C. Calfee & P. Perfumo (Eds.), *Writing portfolios in the classroom* (pp. 203–215). Mahwah, NJ: Lawrence Erlbaum.
Wolf, D., Bixby, J., Glenn, J., & Gardener, H. (1991). To use their minds well: Investigating new forms of student assessment. *Review of Research in Education, 17,* 31–74.

APPENDIX A

Portfolio Structure

Note. The material in this appendix is reproduced with permission (MIIS Graduate School of Language and Educational Linguistics, 2002).

A. Personal Statements

A1. The Portfolio Commentary

The Commentary is the overall preface introducing the Portfolio and its significance. The commentary discusses the criteria for selection of items and explains what the author has learned from being engaged in the Portfolio process. The Portfolio writer may also choose to include a discussion of themes within the Portfolio, but is not required to do so.

A2. The Position Paper

The Position Paper presents the author's current stance as a language professional. It covers the major principles, theoretical positions, and research findings with respect to *language, language learning, and language instruction* that the author finds coherent and compelling. It further explains the implications of these principles in terms of practice (pedagogy, syllabus and materials development, lesson planning, assessment, teacher training).

B. Academic Products

This section contains three items, all based firmly on work undertaken in the various courses in the program.

B1. Revised Projects

The Portfolio will include two papers or projects originally completed as class assignments that have been significantly reworked (i.e., revised, expanded, or both).

B2. Professional Bibliography

The Portfolio will include a bibliography of 25 to 30 articles and books that are deemed significant and useful to the author. Although the bibliography is not annotated, the cover note of about two pages should explain the overall rationale for the bibliography choices and organization; the commentaries elaborate on the value or contribution of particular items.

B3. Summaries of Papers and Projects

The Portfolio also will include 5 to 10 summaries of the candidate's major papers and projects not reworked in Section B1 and that the author considers to be significant. Each summary should contain: (1) a brief account of the circumstances and motivation of the original piece of work (including a heading that notes the course and instructor for whom the assignment was prepared, as well as the date); (2) a summary of the major points or findings and their implications; and (3) a brief mention of the contribution to the author's learning.

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

C. Personal, Practical, and Professional Products

Section C allows the author to present a more personal and introspective account of work and learning, together with actual products and artifacts. The three themes—personal, practical, and professional—should be represented as fully as possible without overwhelming the reader with excessive amounts of material.

C1. Personal and Other Writing

The Portfolio may contain extracts or summaries of journal writing, written observational data, autobiographical writing, correspondence, or any form of writing related to learning or teaching activities.

C2. Professional Products

The Portfolio must contain lesson plans, test items, original or adapted materials, research instruments, syllabus designs, or any other practical product related to teaching and learning. The total number of items may not exceed six.

C3. Audio-Visual Materials; Regalia

The Portfolio may contain any items that help to explicate the author's learning in the Program, the author's approach to teaching, the author's experiences and background relevant to teaching and learning, and so on.

C4. Professional Credentials and Development

The Portfolio must contain the author's CV, plus materials directly related to and documenting in more detail the author's professional experience and contributions.

D. Critiques

This short section contains formal presentations of the three assessment sources:

D1. Self-Assessment

The author presents a two-page critical reflection on the Portfolio, discussing its strengths and weaknesses.

D2. Peer Critiques

Feedback to peers and colleagues is an essential component of a teacher's professional life. This section requires writers to practice and demonstrate the process in two ways: by providing evidence of having given, received, and used meaningful feedback.

D3. Faculty Critique

As part of the review process, at least two faculty members will provide written feedback to the Portfolio after its submission.

APPENDIX B

Assessment Criteria

Note. The material in this appendix is reproduced with permission (MIIS Graduate School of Language and Educational Linguistics, 2002).

1. Depth

This criterion includes two components:

(a) Critical synthesis: At multiple points, the Portfolio must go beyond simply summarizing the ideas of others and must offer evidence of critical thinking and synthesis, combining knowledge with original interpretation.

(b) Coverage of literature: The Portfolio as a whole must convey that the author is accurately acquainted with key thinkers, researchers, and practitioners in the various specialties of Applied Linguistics.

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions

2. Accuracy

Information presented throughout the Portfolio (names, dates, facts, details) is correct. Interpretations, inferences, and conclusions made are fully supported by the evidence and arguments presented. All reference lists are complete and accurate.

3. Contribution to Personal and Professional Development

The Portfolio must represent how the author has grown and progressed both as an intellectual and feeling individual (meeting and reflecting upon new ideas, developing confidence and added strengths through struggle, perhaps initial failure, and subsequent achievement of individual goals) and as a language teaching professional (developing new insights, techniques, and capabilities). Item A1 and all cover notes should unequivocally reflect such growth and development.

4. Voice

The Portfolio, both in its parts and as a whole, must be such that, whether the writing is introspective or academic in register, it is clear that it could only have been written by that particular author. Especially apparent might be aspects of the writer's life experience, academic background, educational philosophy, special talents, sense of humor, and personality.

5. Breadth

The Portfolio must represent the breadth of the program, so that each of the following areas is significantly represented: (1) Language (structure, contrastive studies, language acquisition); (2) Methodology/pedagogy and curriculum; (3) Educational research and assessment; (4) Social and psychological factors in language learning and language use.

6. Integration

The Portfolio must represent a blending of theory, research and practice such that theory and research are explored in terms of practice; practice is justified in terms of theory and research and, in general, the connections among all three are made explicit and richly exemplified.

7. Contribution to the Profession

In at least one section, the Portfolio must represent an actual or potential contribution to the language teaching profession, whether through a novel theoretical insight, an original research study, a useful replication of an existing research study, a new teaching idea, freshly prepared or adapted materials or assessment items.

8. Writing and Presentation

All parts of the Portfolio must be present. The copies must be clear, clean, and of reproducible quality. All Portfolio documents must be free of typographical errors, grammatical inaccuracies, and stylistic infelicities. The writing must be clear, coherent, and well-developed, with a sufficient number of appropriate and effective examples and with consistent formats and referencing procedures.

This content downloaded from 12.233.203.154 on Tue, 2 Dec 2014 11:12:55 AM
All use subject to JSTOR Terms and Conditions