

# **A conceptual model of assessing teaching performance and intellectual development of teacher candidates: a pilot study in the US**

**Kim Hyunsook Song\***

*University of Missouri, St Louis, USA*

A conceptual model was developed in an urban undergraduate education institution in the US to assess teaching performance and intellectual development levels of teacher candidates. Danielson's framework of teaching performance and Perry's pattern of intellectual and ethical development were used for developing a conceptual model. A pilot study was conducted to 282 urban teacher candidates based on the conceptual model to assess their teaching performance and the intellectual development levels. The results from Danielson's and Perry's rubrics showed that the higher students moved to in their teacher education program, the more proficient their teaching performance became and the higher intellectual and ethical development became. Within each of the teaching performance domains and of the intellectual patterns, individual competency patterns varied. This conceptual model may hold promise as a measure of assessing teacher candidates' growth through their performance-based artifacts rather than through the test scores.

## **Introduction**

Traditionally, evaluation of teacher candidates who want to get teaching licensure in public schools in the US of America includes their transcripts, scores on state or national teacher examinations, and letters of recommendation. These traditional evaluation criteria may provide glimpses into the likelihood of being a good teacher, but they do not address teacher candidates' knowledge of content, skills in teaching performance, and/or their intellectual and ethical readiness in order to be successful for their entry level teaching. This study presents a conceptual model that assesses teacher candidates' performance and their intellectual readiness. This model evaluates learning activities (i.e., artifacts) using performance-based standards, and reflective writings using intellectual developmental patterns.

Standards-based assessment reforms attempt to focus on teaching performances of teacher candidates that represent a common conception of instruction to meet the

\*University of Missouri, St. Louis, Division of Teaching and Learning, 369 Marillac Hall, 1 University Boulevard, St. Louis, MO 63121-4499, USA. Email: songk@umsl.edu

needs of diverse learners (Danielson & McGreal, 2000). Proponents of standards-based evaluation suggest that a common dialogue about instruction (Danielson & McGreal, 2000) can be generated when standards and rubrics are used in conjunction with multiple data sources, e.g., unit plans assessed by the standards-based rubric. The assessment data can be used to reinforce and monitor the progress and the weaknesses of teacher candidates' teaching performance. It is believed that such dialogue strengthens both evaluative feedback and objectivity (Kimball, 2002).

Besides examining common conceptions of instruction (e.g., the standards that address content knowledge and pedagogy), teacher education programs need to also examine their students' intellectual and ethical domains (Evans *et al.*, 1998) to determine if they are ready for challenging entry level teaching. In many cases, assessment models of teacher education programs have ignored the importance of evaluating their students' intellectual maturity levels for becoming effective teachers. Preparing morally and intellectually responsible teachers in a diverse society can be one of the most important missions of teacher education programs (Goodlad, 1990). There have been studies that have examined readiness to learn for K-12 school children based on Piaget, Kohlberg, and Chickering, but there has not been much study/evaluation of the moral and intellectual readiness of teacher candidates as part of an evaluation model.

In this study, multiple types of artifacts submitted by teacher candidates are used to assess teaching performance using Danielson's framework of teaching performance, and students' reflective writings are used to examine their intellectual and ethical development level using Perry's intellectual and ethical development pattern.

#### Description of framework for teaching performance

In order to develop a conceptual framework for assessing teacher performance, finding credible, reliable, and valid measures or standards is a critical issue in documenting teacher candidates' knowledge, skills, and dispositions. These standards are the result of an established corpus of knowledge about effective teaching (Ingersoll & Kinman, 2002). One notable development in standards-based assessment is the growing use of Danielson's (1996) *Enhancing professional practice: a framework for teaching* (Kimball, 2002). The Danielson's model serves to illustrate and systematize the responsibilities of teachers to maximize student learning. The framework was designed to evaluate the performance of teachers at different stages and ability levels using explicit standards, multiple data sources and detailed assessment rubrics (Danielson, 1996; Kimball, 2002).

The framework proposed by Danielson organizes 22 components of teaching and 66 elements into four domains of teaching responsibilities. Each component helps to further clarify what is needed to insure competence in a particular domain. Danielson's Domain 1, 'Planning and preparation', incorporates six components that thoroughly embrace the knowledge and skills needed to plan and orchestrate a day of learning in the life of a child regardless of interests, cultural heritage and

available resources. Domain 2, 'Classroom environment', covers five crucial components as varied as providing a physically safe environment to the teacher's expectation of learning and achieving for all students. Domain 3, 'Instruction', includes five components that evaluate teachers' ability to engage students in active learning that promotes critical thinking, decision-making and problem solving while using a variety of instructional strategies to keep students interested in learning. 'Professional responsibilities' outlined in Domain 4 with six components are critical to teacher education programs as the participants are helped to develop professionally and demonstrate a high level of professionalism in their work with children, families, colleagues and the school community (Danielson, 1996; Kimball, 2002).

Danielson's four domains conform with the Interstate New Teacher Assessment and Support Consortium (INTASC) standards that were developed for beginning teachers in the US in 1992 (INTASC, 1992) (see Table 1).

#### Intellectual development

During the 1950s and early 1960s, while serving as the director of Harvard's Bureau of Study Counsel, William G. Perry Jr., along with his associates, engaged in research examining how college students interpret and make meaning of the teaching and learning process (Evans *et al.*, 1998). Perry (1968) acknowledged indebtedness to Piaget and other developmental psychologists such as Kohlberg for his accomplishment

Table 1. Danielson's teaching domains and \*INTASC standards

Danielson's domains and components	INTASC standards
Domain 1: 'Planning and preparation'	1. Content pedagogy
Domain 1: 'Planning and preparation'	2. Student development
Domain 1: 'Planning and preparation'	3. Diverse learners
Domain 2: 'Classroom environment'	4. Multiple instructional strategies
Domain 1: 'Planning and preparation'	5. Motivation and management
Domain 2: 'Classroom environment'	6. Communication and technology
Domain 3: 'Instruction' Domain 4: 'Professional responsibilities'	7. Planning
Domain 1: 'Planning and preparation'	8. Assessment
Domain 3: 'Instruction' Domain 4: 'Professional responsibilities'	9. Reflective practice and professional growth
Domain 4: 'Professional responsibilities'	10. School and community involvement

\*INTASC stands for Interstate New Teacher Assessment and Support Consortium and there are 10 principles for what novice teachers are able to do and need to know before they enter their teaching career in the US.

in the area of intellectual and ethical development of college students. The works of Perry (1968, 1999) hold much explanatory power in suggesting how college students reflect on the information, theories, experiences and opinions that confront them in college classrooms (Bartaglini & Schenkat, 1987).

Perry claims that college students 'journey' through nine 'positions' with respect to intellectual and ethical development. The nine positions are (1) basic duality and (2) full dualism under dualism; (3) early multiplicity and (4) late multiplicity under multiplicity; (5) contextual relativism and (6) pre-commitment under relativism; (7) commitment, (8) challenge to commitment and (9) ongoing commitment under commitment (Rapaport, 2001). These stages can be characterized in terms of the student's attitude towards knowledge. Although Perry's original work identified nine positions, these nine positions are grouped into four major categories to make Perry's scheme more manageable and accessible. 'Dualism', 'Multiplicity', 'Relativism' and 'Commitment' are the four evaluative categories (Winston *et al.*, 1988; Rapaport, 2001).

According to Perry (1968, 1999), the process of college students' intellectual and ethical development is viewed as a progression from dualistic thought to multiplicity, to relativism and finally to stages of commitment, in which the individual uses an internal hierarchy of values to make decisions. This 'journey' is sometimes repeated; and one can be at different stages at the same time with respect to different subjects (Perry, 1981).

Dualism is characterized by dichotomous thinking: the world falls into categories of right/wrong, good/bad and either/or. 'Multiplicity' is the stage where the student comes to see knowledge or truth as being subject to a multitude of interpretations, all having approximately equal weight or validity. 'Relativism' is the stage in which students move from simply acknowledging different perspectives to seeing the various points of view as fitting into a larger whole (King, 1978). 'Commitment' is the stage in which students accept relativism but are able to make 'an active affirmation of themselves and their responsibilities in a pluralistic world' (King, 1978, p. 39). These four categories demonstrate different developmental levels of student thinking (King, 1978; Kirchner & King, 1981).

In the original study, Perry (1968) and his colleagues conducted open-ended interviews with a question like, 'Would you like to say what has stood out for you during the year?' (p. 7), and searched for specific examples within the answers given to identify the student's specific development status. There have been several instruments developed and used to measure college students' development levels based on Perry's theory (Knefelkamp, 1974; Widick, 1975; Baxter Magolda & Porterfield, 1985; Moore, 1989). Whatever assessment instrument or method was used, the goal was to understand college students' perspectives and to form some tentative impressions of where students might be in terms of their intellectual and ethical development (Evans *et al.*, 1998).

## Objectives

The first objective of this study was to describe a professional development process of how the faculty of an urban teacher education program (TEP) located in the US designed a conceptual assessment model. The second objective was to compare and contrast teaching performances and intellectual and ethical development levels among three groups of teacher candidates, i.e., a sophomore pre-professional group (Cohort 1), a junior novice professional group (Cohort 2), and a senior pre-student teacher group (Cohort 3).

A pilot study was conducted to investigate the second objective using the conceptual model. Two research questions were formed for the pilot study:

- Question 1: Will the artifacts of a senior pre-student teacher group (Cohort 3) receive higher scores on a Danielson's teaching performance rubric than those of a sophomore pre-professional group (Cohort 1) and a junior novice professional group (Cohorts 2)?
- Question 2: Will the artifacts of a senior pre-student teacher group (Cohort 3) receive higher scores on Perry's intellectual and ethical development rubric than those of a sophomore pre-professional group (Cohort 1) and a junior novice professional group (Cohorts 2)?

## Methods

### Subjects

Two hundred and eighty-two pre-service students in an urban TEP were the subjects for this study: 95 for Cohort 1, 85 for Cohort 2 and 102 for Cohort 3. A pilot study was done during the fall 2001 semester and the spring 2002 semester at a historically Black college located in a mid-western metropolitan city in the US. Two hundred and twelve subjects (75%) were African-American, 56 (20%) Caucasian and 14 (5%) Hispanic, Asian and/or Native-American. Two hundred and fifty-four (90%) of them were female.

### Procedure

The first objective was to describe how the faculty of an urban teacher education program designed a conceptual assessment model. In this section, a step-by-step procedure is described to meet the objective.

### Preliminary planning

A Danielson's framework and Perry's pattern or framework were selected to examine teaching performance and intellectual development of teacher candidates. A conceptual assessment model was developed based on these two frameworks during preliminary planning meetings with the TEP faculty (see Figure 1). The preliminary

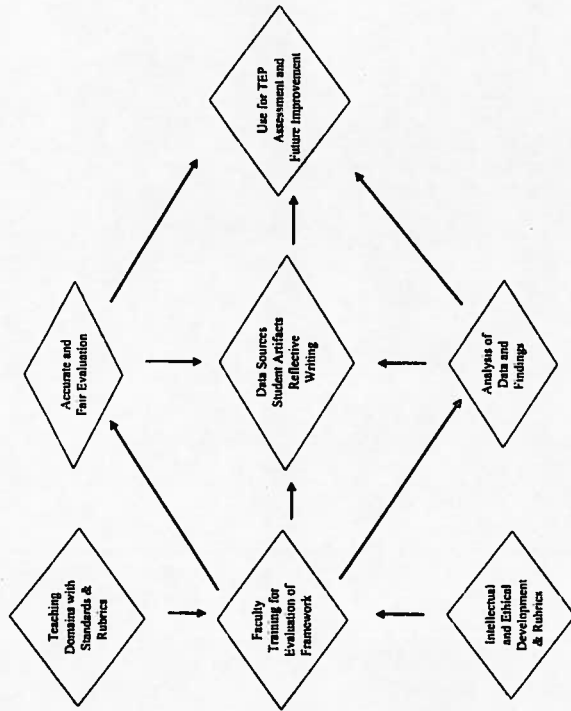


Figure 1. Conceptual model of teaching performance and intellectual development of teacher candidates

planning process began six months before the fall 2001 semester started, and the model was still undergoing revision during and after the fall 2001 semester.

*Development of conceptual model*

*Stage 1. Identification of cohort groups and sample artifacts for each cohort group.* Three cohort groups were identified in Fall 2001: students who had not yet been admitted to the TEP (Cohort 1, pre-professional group), students admitted to the TEP (Cohort 2, novice professional group), and students prior to student teaching (Cohort 3, pre-student teacher group). The TEP faculty identified types of artifacts in Danielson's domains, which could be used for each cohort. For example, in Domain 2, 'Management of the learning environment', types of artifacts included:

- a summary of classroom management theories and models for Cohort 1;
- an interview write-up of students for Cohort 2; and
- a copy of the behavior checklist and classroom rules with consequences for Cohort 3.

Table 2 also includes sample artifacts for Danielson's Domain 1, 'Planning and preparation'.

Table 2. Artifact example in each cohort in Danielson's Domain 1, 'Planning and preparation'

Danielson's domain and INTASC	Cohort 1 (Pre-professional students)	Cohort 2 (Novice professional students)	Cohort 3 (Pre-student teachers)
1. 'Planning and preparation' INTASC: 1, 2, 5, 7, 8	Observation log Summary of theories and models of effective instruction Journals of tutoring Professional journal reviews	Single lesson plans Cooperating teacher's observation forms Instructor's observation forms Reflection journals	Integrated unit plan Lesson enhancer (hands-on and minds-on activities), resources Video tape of teaching with reflection Student work samples with assessment criteria

*Stage 2. Collection of artifacts and reflective writing.* The initial phase of the data collection required TEP students to produce the artifacts as required in their individual course assignments. Next, they selected artifacts from each of their courses that represented their best evidence of content competence. A standards-based performance assessment form was given to the subjects to provide the following information: (1) name of the artifact; (2) the area(s) of study; (3) the name of the course; (4) the certification standard numbers that each of the artifacts addresses, i.e., INTASC (see Table 1); and (5) reflective writings on how each artifact addressed the identified certification standards and how the artifact helped them become more effective in teaching K-12 students in their content areas.

*Stage 3. Artifact assessment using the rubric of Danielson's framework.* Based on the data-driven evidences, the teacher education faculty assessed teacher candidates' artifacts using the rubric adapted from Danielson's (1996) framework for teaching. The quality of the artifacts was scored numerically:

- 1 = Unsatisfactory.
- 2 = Basic.
- 3 = Proficient.
- 4 = Distinguished (see Table 2).

Descriptions of the four levels of performance are included for each element. Those who are at the 'Unsatisfactory' level on a certain element do not meet the minimum performance expectations. Those at the 'Basic' level appear to understand the teaching concept contained in the component, but implementation of the concept is inconsistent or unsuccessful. Those at the 'Proficient' level clearly understand the concept of the element and implement it well. Finally, those at the 'Distinguished' level are highly accomplished teachers who master the element (Danielson, 1996).

Table 3 shows an example of a scoring rubric of 'focus on objectives', which is one of the elements in Domain 1, 'Planning', with the criteria for the four levels. By

establishing a score for each element, it is conceivable for the evaluators to aggregate to a score for each component and then to each domain, and ultimately reach judgments about teaching performance pursuant to the evaluation system (Danielson, 1996; Kimball, 2002).

The following sample artifact, (i.e., objectives of a lesson plan) was rated as 3—'Proficient' based on the rubric (see Table 3). The objectives were written in terms of student learning outcomes, (e.g., '... list people ... and goods ...'). The candidate demonstrated that she could plan lessons developmentally appropriate to the content for the first graders. However, the objectives were still hard to assess (e.g., '... write or draw one person ...'). How was the writing or drawing going to be assessed? The candidate did not provide the criteria for the assessment (see Table 3).

After reading the book *The big green pocketbook*, the students will be able to list people, that the mother and daughter saw, and the goods or services those people provided.

Given the good and services sheet, the students will be able to write or draw one person they know who provides a good and one person they know who provides a service.

**Stage 4. Assessment of writing using the rubric of Perry's intellectual development.** The TEP decided to use Perry's four categories of intellectual and ethical development to examine subjects' development levels as they moved through the curriculum from their inception into their own program to their graduation. A rubric was developed based on the characteristics of Perry's four categories to assess the teacher candidates' intellectual and ethical development (see Table 3). Perry's intellectual development research was based on oral interviews of college students. In this study, however, instead of asking open-ended questions, the participants' reflective writing samples were used to identify their specific developmental status based on the rubric. In December 2001 the reflective statements were collected and in February 2002, eight TEP faculty members and two arts and science faculty rated subjects' reflective writings. Reflection statements were assessed to determine the intellectual development level of teacher candidates using the Perry's rubric numerically:

- 1 = Dualism;
- 2 = Multiplicity.
- 3 = Relativism.
- 4 = Commitment (see Table 4).

The following reflective writing sample was done by a Cohort 3 (pre-student teaching group) and rated as 2.1, which belongs to the range of 'Multiplicity'. The reflection demonstrated the knowledge, skills, and attitudes of what effective teachers need to include in lesson activities. However, there is no description of evidence (e.g., assessment data) or activities why she/he 'feels' that the artifact 'allows her to enhance learning'. The reflection is rather subjective and personal (see Table 4).

This artifact allows me to enhance learning with a new topic on bartering. I created learning experiences that are relevant to learners, and encourage exploration

Table 3. Danielson's rubric example of knowledge of content in Domain 1, Component 1a

Standards MOSTEP: 4 INTASC: 7	Element	B: Focus on Objectives
Domain 1: 'Planning and preparation' Component 1a: 'Demonstrating knowledge of content and pedagogy'	Un satisfactory (1)	Objectives/outcomes do not drive the lesson. (The lesson is a series of activities and not a cohesively planned lesson.) They are not written in terms of measurable student learning.
Basic (2)	Objectives/ outcomes do not drive the lesson, i.e., the lesson. They are written in terms of student learning but may difficult to assess. Candidate designs and plans lessons thoughtfully and appropriate to the content. They are not written as student learning but rather as teacher objectives/goals/activities. A standard lesson design format is used.	Objectives/ outcomes drive the lesson. They are clear, written in terms of student learning and are assessable. Candidate designs and plans lessons thoughtfully and appropriate to the content. Candidate begins to incorporate long-range objectives/goals/activities. A standard lesson design format is used.
Proficient (3)	Objectives/ outcomes drive the lesson. They are written in terms of student learning but may difficult to assess. Candidate designs and plans lessons thoughtfully and appropriate to the content. Candidate begins to incorporate long-range objectives/goals/activities. A standard lesson design format is used.	Objectives/ outcomes drive the lesson. They are clear, written in terms of student learning and are assessable. Candidate designs and plans lessons thoughtfully and appropriate to the content. Candidate begins to incorporate long-range objectives/goals/activities. A standard lesson design format is used.
Distinguished (4)	Objectives/ outcomes drive the lesson. They are written in terms of student learning and are assessable. Candidate designs and plans lessons thoughtfully and appropriate to the content. Candidate begins to incorporate long-range objectives/goals/activities. A standard lesson design format is used.	Objectives/ outcomes drive the lesson. They are clear, written in terms of student learning and are assessable. Candidate designs and plans lessons thoughtfully and appropriate to the content. Candidate begins to incorporate long-range objectives/goals/activities. A standard lesson design format is used.

Source: Adopted from Danielson (1996).

Table 4. Perry's rubric example of 'Dualism'

Level	Description of each level
'1' for 'Dualism'	<p>Knowledge is quantitative. Meaning is divided into two realms: black/white, right/wrong. Questions have immutable, objective and right answers. Certainty yields to uncertainty and ambiguity. Students generally believe authorities possess valuable wisdom that contains eternal truths (e.g., all problems are solvable. Student's task is to learn the right solutions and ignore the others.)</p>

Source: Developed by the researcher based on Perry (1968, 1999).

and problem solving. . . . This lesson activity helped me become a more effective teacher in the following ways. I feel that I was able to relate to the children. . . . I was able to experience the hands-on with my students and be involved with the process of learning. . . . I feel this activity helped elementary students to learn better. . . . I feel that my students enjoyed the lesson and had fun while they learned.

### Result of pilot study

The result of the pilot study was to examine the second objective, 'to compare and contrast teaching performances and intellectual and ethical development levels among three groups/cohorts of teacher candidates'. The data reported in this pilot study used mean analysis for each variable (domain) by cohorts (teacher candidate classifications).

**Question 1:** Will the artifacts of a pre-student teacher group (Cohort 3) receive higher scores on a Danielson's teaching performance rubric than those of a pre-professional group (Cohort 1) and a novice professional group (Cohorts 2)?

**Result 1:** Table 5 illustrates with mean scores and standard deviations in critical components of each of the four domains. In Domain 1, 'Planning and preparation', the novice professional student group (Cohort 2) and the pre-student teacher group (Cohort 3) were rated 'Proficient' in each of the four components of the domain, i.e., 'Knowledge of subject', 'Focus on objectives', 'Knowledge of how students learn' and 'Assessment of objectives' (mean range = 3.14-3.50), whereas the pre-professional student group (Cohort 1) demonstrated 'Basic' competencies (mean range = 2.57-2.78) across the domain. Interestingly, the novice professional group (Cohort 2, mean = 3.50) was rated by the faculty consistently higher on a component of 'Assessment of objectives' than the pre-student teacher group (Cohort 3, mean = 3.26) (see Table 5).

In Domain 2, 'The classroom environment', the participants were rated 'Proficient' across the three cohort groups (range = 3.1-3.31) in two components, i.e., 'Classroom management' and 'Discipline'. The pre-student teacher group (Cohort 3, mean = 3.31) showed the highest mean in this domain. The novice

Curriculum cohorts		Domain 1: 'Planning and preparation'		Domain 2: 'Classroom environment'		Domain 3: 'Instruction'		Domain 4: 'Professional responsibilities'		Communication		Perspectives of self and others								
Critical components		Knowledge of subject		Focus on objectives		Knowledge of how students learn		Assessment of objectives		Critical components		Instructional strategies		Critical thinking		Communication		Perspectives of self and others		
Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3
Mean	2.78	2.78	2.70	2.69	2.57	2.78	2.78	2.78	2.78	2.78	2.78	2.78	2.78	2.78	2.78	2.78	2.78	2.78	2.78	2.78
SD	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
Mean	3.14	3.14	3.20	3.06	3.10	3.14	3.14	3.14	3.14	3.14	3.14	3.14	3.14	3.14	3.14	3.14	3.14	3.14	3.14	3.14
SD	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83
Mean	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51
SD	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71
Mean	3.30	3.30	3.30	3.30	3.30	3.30	3.30	3.30	3.30	3.30	3.30	3.30	3.30	3.30	3.30	3.30	3.30	3.30	3.30	3.30
SD	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70
Mean	3.26	3.26	3.26	3.26	3.26	3.26	3.26	3.26	3.26	3.26	3.26	3.26	3.26	3.26	3.26	3.26	3.26	3.26	3.26	3.26
SD	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60

Table 5. Mean scores and standard deviations of teacher candidates' performance

Note. 'Unsatisfactory' = 1; 'Basic' = 2; 'Proficient' = 3; and 'Distinguished' = 4

professional student group (Cohort 2, mean = 3.13) scored lower on a component of 'Discipline' than did the pre-professional student group (Cohort 1, mean = 3.28) (Table 5). In Domain 3, 'Instruction', the participants were rated as 'Proficient' across the three cohort groups on two components, i.e., 'Instructional strategies' and 'Critical thinking' (range = 3.15-3.35) (see Table 5).

In Domain 4, 'Professional responsibilities', the participants were rated by the faculty as 'Proficient' across the three cohort groups on three components of the domain (range 3.01-3.37), i.e., 'Communication', 'Commitment' and 'Perspectives of self and others'. The pre-professional group was scored by the faculty as 'Basic' in 'Communication' (mean = 2.94) but 'Proficient' in 'Commitment' (mean = 3.02) and 'Perspectives of self and others' (mean = 3.16). The novice professional student group and the pre-student teacher group were rated highest in 'Communication' (mean = 3.18 and 3.37) followed by 'Commitment' (mean = 3.14 and 3.37) and then 'Perspectives of self and others' (mean = 3.13 and 3.30) (see Table 5).

**Question 2:** Will the artifacts of a pre-student teacher group (Cohort 3) receive higher scores on a Perry's intellectual and ethical development rubric than those of a pre-professional group (Cohort 1) and a novice professional group (Cohorts 2)?

**Result 2:** Table 6 illustrates that the pre-professional student group (Cohort 1) was scored at an upper level of 'Dualism' (mean = 1.85); the novice professional student group (Cohort 2) was also scored at the level of 'Dualism' (mean = 1.72), with a mean score lower than that of the pre-professional group (Cohort 1). The pre-student teacher group (Cohort 3) scored at a level of 'Multiplicity' (mean = 2.16). No cohort group scored at the stage of Perry's 'Relativism' or 'Commitment' (see Table 6).

### Discussion

A small urban undergraduate teacher education program designed a conceptual model of assessing teacher candidates' teaching competencies and their intellectual and ethical development. The act of designing a conceptual model of assessment that

Table 6. Mean scores and standard deviations of teacher candidates' intellectual development

Curriculum cohorts		Mean	
Cohort 1	Mean	1.85	
	SD	1.33	
Cohort 2	Mean	1.73	
	SD	.83	
Cohort 3	Mean	2.16	
	SD	.99	

*Note.* 'Dualism' = 1; 'Multiplicity' = 2; 'Relativism' = 3; and 'Commitment' = 4.

incorporates teaching proficiencies and developmental levels requires a great deal of time and effort. Unlike previous assessment systems, which relied on licensure examination, Praxis II, and transcripts, the conceptual model established more structures, increased opportunities for dialog among the TEP faculty and between the faculty and students, and provided multiple data sources for assessing teaching performances and intellectual readiness of the TEP students. The study was done based on the evaluation of the teacher candidates' artifacts and their reflective writings. The assessment process provided practical sources for diagnosing the strengths and weaknesses of teaching performance and intellectual and ethical preparedness.

In the pilot study, the results from Danielson's scoring rubric on teaching practice showed in general that the higher cohort group (Cohort 3) acquired higher scores than the other two groups in four domains even though individual competency patterns varied within each of the four domains. The scores from Perry's rubric on intellectual and ethical development also illustrated that the higher cohort group (Cohort 3) had higher scores than the other groups.

The scores from Danielson's rubric showed that pre-student teachers reached the level of 'Proficient'. When they student teach, they may develop their teaching skills and knowledge to reach the level of 'Distinguished' even though this study has no data to support this assumption. The result, however, showed an unexpected pattern between Cohort 1 and Cohort 2. The novice professional student group's (Cohort 2) artifacts were consistently rated lower than the pre-professional student group's (Cohort 1) throughout Danielson's domains. This was unexpected in that one would think the juniors (Cohort 2) would have had richer and more varied educational experiences than the sophomores (Cohort 1). This anomaly might be explained in several ways. Introductory teaching methods courses often focus on those basic skills and strategies that lay the foundation for higher level teaching skills and strategies. Perhaps Cohort 1 students, having little actual teaching experience, saw teaching as inflexible, not varying at all from what was taught as foundational theory. Cohort 2 students who have had some classroom observation and tutoring experiences possibly realized they lacked the required knowledge and skills in the critical teaching component, e.g., classroom management and assessment and discovered that real practice is different from what they learned in foundation classes. In addition, courses may not have been taken in the designated order. As we assessed cohort membership, we found that although students were assigned according to program classification they may not have had all of the curriculum necessary to that classification. Individual differences in students and in their approach to learning may have affected the result. The subjectivity of the raters may be another influencing factor. Since this was the first time for faculty members to use Danielson's rubric, it is possible that they did not have enough practice to be expert evaluators.

Students in Cohorts 1 and 2 received a rating of 'Dualism' and those in Cohort 3 received an overall rating of 'Multiplicity' on Perry's framework. The intellectual and ethical lens of Cohort 1 and Cohort 2 students appears to be narrowly focused

regarding what is required of teachers. As the students progress through their general education coursework, their lens widens. Cohort 3 students, rated in the 'Multiplicity' area, may have realized that the world is more than black and white and moved toward more choices. The participants, however, did not reach the next levels, 'Relativism' and 'Commitment'. Kitchener and King (1981) found in their study that most traditional-aged college students did not progress beyond the 'Multiplicity' stage during their undergraduate years. Thus, the results from the pre-service teachers fit into the average college students' cognitive developmental level, 'Multiplicity'.

Perry's reflection rating showed an unexpected pattern between Cohort 1 and Cohort 2. The pre-professional student group (Cohort 1) received a higher rating than the novice professional student group (Cohort 2) even though both groups belonged to the level of 'Dualism'. The students in Cohort 1 usually come in with 'cut and dry' solutions whereas those in Cohort 2 have begun to see solutions more in 'shades of gray' realizing that there may be situations that require more than a 'cut and dry' solution.

As Ingersoll and Kinman (2002) argued, individual education institutions must translate their exit assessment criteria into measurable benchmarks for novice teachers. The character of that translation will reflect the underlying conceptual framework within which the institution operates (NCATE, 2001). This study described the process of converting their assessment from test-score based evaluation to performance-based evaluation using students' artifacts and reflective writing. These artifacts and writings were assessed by a conceptual assessment model, which provides measurable benchmarks (Kemp *et al.*, 2003). Undergraduate education students may enter their preparation program with preconceived notions about teaching and their competence to enter the profession. Those perceptions may or may not be consonant with the aims of the institution. This conceptual model with outcome-based criteria provides specific constructs that undergraduate education students perceived to underlie their entry teaching. The teacher candidates may accommodate their beliefs about teaching to conform to the realities they experience in coursework and practical experience by reviewing their weaknesses and strengths with assessment criteria, (e.g., criteria of Danielson's Component 1b: 'Focus on objectives' in Table 3 and Perry's criteria of 'Dualism' in Table 4).

In summary, the results from Danielson's and Perry's rubrics showed that the higher cohort the subjects moved to in their teacher education program, the more proficient their teaching performance became and the higher their intellectual and ethical development became. None of the artifacts submitted by teacher candidate reached Danielson's 'Distinguished' level for teaching performance, nor Perry's 'Commitment' level for intellectual development. This conceptual model, however, may hold promise as a measure of assessing teacher candidates' growth through their performance-based artifacts rather than through their test scores.

The following are suggested for implementing a more effective assessment of teacher candidates' performance and of their intellectual growth:

- In the earliest stage of a teacher preparation program, the conceptual model for assessing teaching performance and intellectual development needs to be introduced to teacher candidates and studied across an array of courses.
- The teacher education program needs to be designed carefully with a balance between content and field experiences in order to promote the successful adoption of the conceptual model (Sharpton *et al.*, 2002).
- A replication study should be done as a longitudinal study, which may include Praxis II scores and K-12 student achievement test scores, to examine the improvement of participants' teaching competencies based on utilizing this conceptual mode.
- A study should be done through observation and interviews of not only teacher candidates, but also cooperating teachers, teacher educators, and public school (K-12) students to document if the performance-based assessment using the conceptual model helps them assess their teaching performance and their intellectual and ethical readiness.

## References

- Battaglini, D. J. & Schenker, R. J. (1987) *Fostering intellectual and ethical development in college students: the Perry and Toulmin models*, ERIC Digests (ED284272) (Urbana, IL, ERIC Clearinghouse).
- Baxter Magolda, M. B. & Porterfield, W. D. (1985) A new approach to assessing intellectual development on Perry scheme, *Journal of College Student Personnel*, 26, 343-351.
- Bizzell, P. (1984) William Perry and liberal education, *College English*, 46, 447-454.
- Danielson, C. (1996) *Enhancing professional practice: a framework for teaching* (Virginia, Association for Supervision and Curriculum Development).
- Danielson, C. & McGreal, T. L. (2000) *Teacher evaluation to enhance professional practice* (Virginia, Association for Supervision and Curriculum Development).
- Dwyer, C. A. (1994) *Development of the knowledge base for the PRAXIS III: classroom performance assessment criteria* (New Jersey, Educational Testing Service).
- Evans, N. J., Forney, D. S. & Guido-DiBrito, F. (1998) *Student development in college: Theory, research, and practice* (San Francisco, Jossey-Bass).
- Goodlad, J. I. (1990) The occupation of teaching in schools, in: J. I. Goodlad, R. Soder & K. A. Sirotnik (Eds) *The moral dimensions of teaching* (San Francisco, Jossey-Bass).
- Ingersoll, G. M. & Kinman, D. (2002) Development of a teacher candidate performance self-assessment instrument, *Action in Teacher Education*, 23, 4-9.
- Interstate New Teacher Assessment and Support Consortium (1992) *Model standards for beginning teacher licensing and development: a resource for state dialogues* (Washington, DC, INTASC).
- Kemp, L., Song, K. & Johnson, P. (2003) *Design and implementation of a performance-based assessment system for the preparation of effective teachers for a diverse society* (Missouri, Harris-Stowe State College).
- Kimball, S. M. (2002) Analysis of feedback, enabling conditions and fairness perceptions of teachers in three school districts with new standards-based evaluation systems, *Journal of Personnel Evaluation in Education*, 16, 241-268.
- King, P. M. (1978) William Perry's theory of intellectual and ethical development, in: L. L. Knefelkamp, C. Widdick & C. A. Parker (Eds) *New directions for student services: applying new developmental findings* (San Francisco, Jossey-Bass).



- Kirchener, K. S. & King, P. M. (1981) Reflective judgment: concepts of justification and their relationship to age and education, *Journal of Applied Developmental Psychology*, 2, 89–116.
- Kniefelkamp, L. L. (1974) *Developmental instruction: fostering intellectual and personal growth in college students*. Unpublished doctoral dissertation, Minneapolis, University of Minnesota.
- Moore, W. S. (1989) The learning environment preferences: exploring the construct validity of an objective measure of the Perry scheme of intellectual development, *Journal of College Student Development*, 30, 504–514.
- NCATE (National Council for the Accreditation of Teacher Education) (2001) *Professional standards for the accreditation of schools, college and departments of education*. Available online at: <http://www.ncate.org/2000/2000stds.pdf> (accessed 12 March 2004).
- Perry, W. Jr. (1968) *Patterns of development in thought and values of students* (Massachusetts, Harvard University Press).
- Perry, W. Jr. (1981) Cognitive and ethical growth: the making of meanings, in: A. W. Chickering & Associates (Eds) *The modern American college: responding to the new realities of diverse students and a changing society* (San Francisco, Jossey-Bass).
- Perry, W. Jr. (1999) *Forms of intellectual and ethical development in college years* (San Francisco, Jossey-Bass).
- Peterson, C. H. (1982) *A century's growth in teacher evaluation in the United States* (New York, Vantage).
- Rapaport, W. J. (2001) *Journey along the 9 Perry positions: William Perry's scheme of intellectual and ethical development*. Available online at: [www.cs.buffalo.edu/~rapaport/perry.positions.html](http://www.cs.buffalo.edu/~rapaport/perry.positions.html) (accessed 5 February 2002).
- Reeves, D. B. (2002) *The Leader's guide to standards: a blueprint for educational equity and excellence* (San Francisco, Jossey-Bass).
- Sharpton, W., Casbergue, R. & Caside, K. (2002) Pre-service preparation for the urban context, *Metropolitan Universities*, 13, 21–32.
- Widick, C. (1975) *An evaluation of developmental instruction in a university setting*. Unpublished doctoral dissertation, Minneapolis, University of Minnesota.
- Winston, R. B. Jr., Bonney, W. C., Miller, T. K. & Dagley, J. C. (1988) *Promoting student development through intentionally structured groups: principles, techniques, and applications* (San Francisco, Jossey-Bass).

## Out of balance: lecturers' perceptions of differential status and rewards in relation to teaching and research

Pat Young\*

University of the West of England, Bristol, UK

One of the few areas of consensus in the literature of higher education concerns the status of teaching. Unanimously, writers report the low status which higher education institutions give to teaching as an activity. This article draws on research investigating activities and perceptions of staff in a single discipline: social policy. The question on rewards for effort in improving teaching provoked more strength of feeling and a greater degree of consensus than any other in the interviews. With similar unanimity to that found in the secondary sources, the respondents perceived teaching to be accorded low status, with rewards of tenure and promotion accruing to research or administration. The article suggests that the impact of new developments to enhance teaching and learning are undermined by the persistent low status accorded to teaching.

### Introduction

This article brings together evidence from the literature on teaching in higher education, and data from a study of lecturers working in a specific discipline in the UK, to argue that a major barrier to developing teaching and learning in higher education is the low status and lack of reward for effort devoted to improving teaching.

The source of primary data is a research study with social policy lecturers, carried out by myself and Dr Zoë Irving, working within the—then Learning and Teaching Support Network, now Higher Education Academy—Subject Centre for Social Policy and Social Work (SWAP). The project was funded by the four higher education councils of the UK and involved a series of in-depth, semi-structured interviews with 46 lecturers working in higher education institutions (HEIs) across the UK. Twenty institutions were randomly selected from a list of all HEIs offering an undergraduate programme with the title of social policy, social administration or public policy. The sample reflected the original population in the balance of 'old' 'pre-1992' institutions to those gaining university status since the end of the binary system of higher education in 1992. It also represented the four countries of the UK.

\*Faculty of Health and Social Care, University of the West of England, Blackberry Hill, Bristol BS16 1DD, UK.