

Statement of Project Focus and Rationale:

The Title II Teacher Quality Enhancement Project aims to address the future educational needs of the State of Connecticut, by investigating ways in which teacher quality can be improved and enhanced. In order to address these needs, the Connecticut State Department of Education awarded a number of grants to research teams at public and private universities throughout the state. An interdepartmental, collaborative team at the University of Hartford was the recipient of one of these grants, and this team focused on the development of a survey instrument to assess teacher candidate dispositions.

Specifically, the University of Hartford team was composed of professors from the Division of Education and the Psychology Department of the College of Arts and Sciences, and K-12 teachers from the University of Hartford Magnet School. Thus, the team combined the strengths of scholars experienced in the areas of education and test development/validation, as well as the knowledge and skills of current educational practitioners.

The goal of the team was to develop a practical, efficient survey to assess candidates' dispositions and to provide feedback about candidates' readiness at various transition points throughout the education program. This survey would allow for the assessment of dispositions identified as important for teaching, and was intended to be completed by candidates at the beginning of the sophomore year (pre-test) and at the end of the senior year (post-test). Administering the survey at these various stages and providing the candidates with feedback would:

- (a) enable advisors to assist candidates in understanding the requisite dispositions for teaching, and in developing strategies to improve their dispositions,
- (b) allow candidates to improve their dispositions and to track changes in dispositions, and
- (c) provide candidates with sufficient self-knowledge to determine whether to continue in the teaching profession.

Thus, the survey and feedback system would give candidates feedback on areas for improvement as well as assessment of improvement, while allowing

candidates to self-select out should they come to recognize that the dispositional requirements of teaching are incompatible with their personal dispositions.

Defined Deliverables

Currently, psychologists and employers have a variety of psychological measures of dispositions to choose from, including emotional intelligence measures and personality tests (e.g., Myers-Briggs, 16 PF, NEO-PI-R) that assess traits such as extraversion, emotional stability, and conscientiousness. However, these tests typically assess global traits, whereas the main focus of the Title II Teacher Quality Enhancement Project at the University of Hartford was to develop a survey that would focus on specific dispositions that are determined to be necessary for effective teaching. Thus, the development of a unique survey targeting these specific dispositions was the goal of the University of Hartford team. This survey (see Section 2) is the primary deliverable from this project.

In addition, it was determined that the survey would eventually be administered online (i.e., on computer) and be scored mechanically using an established scoring rubric. Given the large number of students that enter the teaching curriculum, hand-scoring of responses was determined to be prohibitively time-consuming, particularly when the survey would be administered a several points throughout the program. Therefore, a multiple-choice or Likert-type response scale was identified as preferable, and the current version of the survey (see Section 2) uses the Likert-type response scale.

A scoring rubric which identifies the linkage between relevant dispositions and survey items is also presented in Section 2.

Survey Development Process

In June of 2002, the University of Hartford team met initially, and team members were assigned to one of two groups (see Appendix A). All members attended the meetings from June 24th to September 24th 2002, with the exception of Peter Oliver and Tony Crespi who indicated that they were not able to attend meetings. Appendix B contains minutes from the initial and subsequent meetings of the groups.

At subsequent meetings, the team determined that the dispositions assessed by the survey would follow the NCATE framework, and the INTASC and other State competencies would also be used to frame the dispositions to be assessed. Specifically, the NCATE standards that were to be assessed included (1) Development, Learning, & Motivation, (2) Curriculum, (3) Instruction, (4) Assessment, and (5) Reflection. Although there was some consideration given to using a pre-existing personality test (i.e., the 16 PF), the consensus was that case studies or “vignettes” would be most appropriate for assessing the requisite dispositions. Specifically, candidates would be required to interpret these vignettes and their interpretations of the vignettes would reflect their dispositions.

Because the survey was intended to be administered on computer, the team determined that short vignettes and short answers would be preferable, to allow for scrolling on the computer screen.

The team considered to possibility that candidates' interpretations of the vignettes could take the form of narrative responses, but in order to facilitate computerized administration and scoring, a multiple-choice or Likert-type scoring system was selected, with four responses for each vignette. Team members were asked to write vignettes to assess the dispositions in the NCATE framework, with the goal to have approximately four or five vignettes for each NCATE standard.

It is relevant to note that such types of assessments are sometimes used by employers and human resource professionals in the process of personnel selection. Specifically, Industrial/Organizational Psychologists have developed similar tests for use in selecting employees for a variety of jobs. Such tests, termed situational judgment tests or inventories (SJTs or SJIs), present a typical work-related scenario and provide several possible actions or responses to that scenario. Test takers typically indicate which response is best or most like them and which is worst or least like them. Researchers have found that such tests typically tap multiple dispositions within an item, given the complexity of the scenarios.

Team members submitted vignettes to the assistant principal investigator, who compiled them into a single document. The team as a whole then reviewed the vignettes for clarity, realistic portrayal of classroom situations, and fit with the targeted NCATE standard. Some vignettes were removed and/or altered, and additional vignettes were added. Of the five NCATE standards, Curriculum and Assessment were identified as lacking a sufficient number of vignettes (i.e., less than four or five) for proper assessment and reliability, and thus more vignettes were written. In contrast, there were more than five vignettes for the Development, Learning, & Motivation and Instruction standards. The team determined to retain all of the vignettes until after the pilot test and to eliminate poorly-functioning vignettes based on the pilot results. Similarly, some vignettes had greater than the targeted four responses, but all responses were retained until after the pilot for the same reasons. A Likert response scale with six levels ranging from 1 = Strongly Disagree to 6 = Strongly Agree was also preferred over the multiple choice scale, in order to maximize variance.

Irv Richardson, the consultant for the Title II project provided the team with feedback on the survey (Appendix C). This feedback was used to improve the survey. Specifically, the dispositions defined by the NCATE standards were clarified, and their linkages to the survey responses were clarified in a matrix. More vignettes were written to assess the dispositions from the Curriculum and Assessment standards. All of the vignettes were also re-examined and in some cases revised to reflect elementary school (i.e., K-5) settings. Finally, the items were revised to be more consistent in format and verb usage.

This revised draft version of the survey along with matrix linking responses to NCATE Standards and dispositions, is provided in Section II. This version was presented at the Title II Teaching Quality Enhancement Workgroup Grants Conference on December 13, 2002.

Further feedback on the survey was provided by Georgette Nemr and Irv Richardson on February 4, 2003. This feedback led to further clarification and refinement of the targeted dispositions and simplification of the survey format. Specifically, two broad dispositions were defined as critical for candidates to possess:

Professionalism

- Engages in ethical practices
- Values the collaboration and involvement with students, peers, families, and other professionals
- Displays initiative and dependability
- Demonstrates commitment to continuous improvement in professional practice

Reflective Practice

- Demonstrates commitment to life-long learning
- Values self-directed learning, critical thinking, and self-understanding
- Believes all students can learn

These dispositions formed the basis for the new survey format, such that two longer vignettes were developed to tap these two dispositions, respectively. These longer vignettes took the form of a case study of a day in the life of an elementary school teacher and were divided into four sections for greater ease of administration and to reduce cognitive load on the candidates. Subsequent to each section are a series of statements about the section with the Likert-format response options as discussed above (i.e., 1 = Strongly Disagree to 6 = Strongly Disagree). In addition, candidates are asked to write their rationales for each response. This information may be used to further refine the survey and may also assist advisors in targeting candidates' dispositional issues.

A scoring rubric was developed by having each of the four core subgroup members independently link the items on a vignette with the dispositional aspects (listed above, under each disposition). Members also indicated whether the item should be reverse-scored (i.e., if the item was negatively worded). These linkages were then compiled and where three or four of the members agreed to the linkage, or no members indicated that a linkage was present, this was considered complete

agreement. However, when one or two members considered a linkage present, consensus was reached through discussion by the subgroup members. This process was intended to provide evidence for the construct validity of the survey, in addition to establishing the basis for the scoring rubric.

The final version of this survey, along with a matrix linking each statement to the pertinent dispositional characteristics and a scoring rubric, are presented in Section II. Discussion of the pilot testing of this survey is presented in Section III.