

A Four-Year Longitudinal Study of Principled Judgment Reasoning in Teacher Education Students

This paper has four objectives. First, to conceptually link studies of principled judgment reasoning to recent efforts by national teacher accrediting bodies to include professional dispositions as a teacher education program goal. Second, the investigator reports four-year longitudinal results of a character development study of teacher education students as measured by principled judgment reasoning scores. Four-year results are compared with other four-year longitudinal samples. Third, the investigator examines correlations between principled judgment reasoning and prospective teachers' performance during student teaching. Finally, implications are drawn for teacher education programs.

Overview and Definitions

Recently, there have been revisions in accreditation standards for teacher education programs and programs for the preparation of principals. These new accreditation standards for colleges of education reflect an interest in designing, implementing, and evaluating educational programming that addresses knowledge, skills, and professional dispositions of future teachers (NCATE, 2000). Changes in the accreditation standards for teacher education reflect multidisciplinary interest in professional dispositions across professional preparation programs in higher education

Yet, there are questions about what is meant by dispositions. Katz and Rath (1985) proposed that professional dispositions be added to professional knowledge and skills as goals for teacher education. They defined dispositions as trends in professional judgment and professional action as contrasted with habits, skills, attitudes, and traits (Katz & Rath, 1985, p.301). For example, a trend in one's professional actions and decision making to employ principles when faced with conflict-laden and ill-structured situations in classroom teaching would represent one example of a professional disposition. If such judgment occurs on most occasions when the teacher is faced with such problems, then the teacher can be said to have a disposition to use such reasoning. While some dispositions would be conscious and deliberative, others are habitual and "automatic."

In contrast to dispositions, a skill connotes some level of mastery in a technique, and an attitude represents a relatively enduring organization of beliefs around an object, predisposing one to respond in a preferential manner. Thus a skill is something one either possesses or not, and an attitude is a pre-disposition to act, rather than a trend in professional judgment and professional action.

In concert with the new NCATE standards, colleges of education are beginning to articulate the dispositions that the faculty value in teachers and other professional school personnel (NCATE, 2000). Further, institutions must provide a conceptual understanding of how knowledge, skills, and dispositions relate to diversity and technology, and to describe how these competencies are integrated across curriculum, instruction, field and clinical experiences, assessments, and evaluations (NCATE, 2000, p.3). Yet, colleges

of education are only now beginning to define, design, and assess their programs in light of the new NCATE standards which include dispositions.

For the purposes of this study, disposition was defined as trends in professional judgment and professional action. A teacher observed to frequently demonstrate empathy in a number of classroom and schooling contexts has a disposition to be empathic. As one can therefore see, disposition serves both descriptive and forecasting functions. In identifying professional dispositions, NCATE is interested in those intentional dispositions that are related to effective teaching. Yet too little research has addressed which dispositions have positive correlations to effective teaching performance.

In this study, the investigators posited three dispositions that discriminate between more and less successful adults and teaching professionals (Sprinthall, Reiman, & Thies-Sprinthall, 1996). They include:

- The teacher as epistemologist and compassionate manager of learning and instruction.
- The teacher as committed representative of democratic principles, and
- The teacher as self-actualized individual.

These three dispositions were operationalized in the following way. The first disposition, teacher as epistemologist would be assessed by measuring students ability to symbolize professional experience and to flexibly and reflectively solve problems. The second disposition, teacher as committed representative of democratic principles, was assessed through measurement of prospective teachers principled reasoning judgments. The third disposition, teacher as self-actualized individual was assessed by measuring students' ability to interpret the views of others, and to be identity resolved. In this paper, we report longitudinal research on the second disposition – to make principled judgments.

Employing Principled Judgments: A Needed Dispositional Goal for Teacher Education

Teachers have considerable influence on students' moral and intellectual development (Chang, 1994; Mentkowski, 2000; Yost, 1997), and educators propose that more attention be given to these teacher dispositions in colleges of education and schools (Goodlad, 1994; Oser, Dick, & Patry, 1992, Strike, 1996). The limited studies of teachers note that moral decision making occurs daily as teachers select learning outcomes, structure classroom activities, allocate professional resources, and engage students, parents, and the community in learning and dialogue (Maslovaty, 2000; Oser & Altof, 1993). As Rest, Narvaez, Bebeau, and Thoma (1999) acknowledge, persons' ability to think abstractly and to flexibly consider moral problems from multiple perspectives is an indication of postconventional principled moral reasoning. Such an ability is one hallmark of wise professional judgment and action. Yet there are few studies that examine the relationship between principled judgment reasoning and effective teaching performance in teacher education programs. It is worth noting as well, that a small number of studies of students majoring in education indicate that education students' principled moral reasoning is lower than that of other college majors (Lampe, 1994; McNeel, 1994). The study is now described.

Method

Participants

Participants included 49 preservice teacher education students (middle school and secondary). Seventy-one percent of the participants were females (n=35) and twenty-six percent of the participants were male (n=13). Gender was not reported by one participant (n=1). Females had an average SAT V of 500, an average SAT Q of 553, and an average total SAT score of 1053. Males had an average SAT V of 518, an average SAT Q of 593, and an average total SAT score of 1111. Females had an average entering academic index 2.75, and a graduating GPA of 3.35, while male participants had an average entering academic index of 2.78 and a final college GPA of 3.14. All entering students were from the state in which the university resided.

Instrumentation: Principled Judgment Reasoning

Students' level of principled judgment reasoning was assessed through the administration of The Defining Issues Test (DIT) (Rest, 1979). The test was recently reviewed by Rest, Narvaez, Bebeau, and Thoma (1999). The test emphasizes cognition and developmental schemas. They postulate three structures of ethical thinking: the personal interest schema, the maintaining norms schema, and the postconventional schema.

Data has been gathered, and only recently interpreted, on over 50,000 persons (Rest, Narvaez, Bebeau, & Thoma, 1999). The construct validity questions and reliability issues have been addressed by confronting a number of issues (Rest, Narvaez, Bebeau, & Thoma, 1999). These items are reviewed as they are important to the present study described in the paper. Among the items, the DIT should: 1) differentiate groups to be of greater or lesser expertise in moral reasoning, 2) show samples making significant upward change in longitudinal studies, 3) demonstrate sensitivity to interventions designed to improve moral reasoning (e.g., show pre-/posttest gains on moral education programs), 4) show evidence of a developmental hierarchy, 5) predict real-life moral behavior, 6) predict political attitudes and the way a person participates in the larger community, and 7) demonstrate adequate reliability. Across over 400 studies by independent researchers, these claims have been vigorously met. The test does differentiate expertise in moral reasoning (Thoma, 1986) and does show samples making significant upward changes across nine longitudinal studies (McNeel, 1994). As well, the test is sensitive to interventions designed to improve moral reasoning (Oja & Reiman, 1998; Reiman, 1999, Rest & Narvaez, 1994; Sprinthall, Reiman, & Thies-Sprinthall, 1996).

As well, links have been made to behavior. Higher scores on the DIT have been connected to various prosocial behaviors including community involvement ($r=.31$) and civic responsibility ($r=.44$) (Thoma, Rest, & Barnett, 1986). As well, review of 60 published studies show significant links between DIT scores and teacher decision making, nurses' clinical performance, auditors detection of fraudulent financial reports, and accountants perceptions of management's integrity (Rest & Narvaez, 1994). Regarding reliability, estimates for 20 years of Cronback alpha for the DIT have been in the high .70s (Rest, Narvaez, Bebeau, & Thoma, 1999). One final note, the field of moral psychology conceptualizes the entire moral domain to include at least four major internal components that lead to ethical behavior: moral sensitivity, moral judgment, moral motivation, and moral character (Rest, 1983).

Student Teaching Performance

Student teaching performance was assessed using a measure of teaching performance that is used by the university's teacher education program. The instrument requires independent raters (university faculty and cooperating teacher) to assess instructional planning and curriculum design, instructional implementation including quantity and quality of engagement with learners, and student teacher success in reaching outcomes. The assessment utilizes a six point scale. A score of 1.0-2.0 suggest that the student teacher demonstrates limited use of instructional practices and exhibits no confidence with a variety of curricular approaches and instructional models and methods. A score of 3.0 - 4.0 suggests that the student teacher exhibits routine use of instructional practices and exhibits emerging confidence with a variety of curricular approaches and instructional models and methods. Scores between 5.0 and 6.0 suggest that the student teacher exhibits integrative use of instructional practices, and shows high levels of confidence and flexibility. No formalized assessments of validity or reliability of this instrument have been conducted. As well, there was no statistical analysis of interrater reliability. Scores from the two raters were averaged for each student in the sample.

Procedures

The teacher education students in the sample were administered a battery of assessments including the DIT during their first week at the University. At that time, students were informed about the purpose of the study, advised that participation was voluntary, and assured that all results would be confidential. Students completed a letter of informed consent to voluntarily participate. The battery of assessments took 60- to 90- minutes to complete. The DIT was the first instrument in the packet to complete. After testing was completed, participants placed all materials into their envelopes. Investigators collected all envelopes and then assigned a four digit number to each set of assessments. Only the principal investigator had access to the names and assigned numbers of the study participants. To protect participants' identities, numbers only were used in data analysis. Posttesting was completed during the month of April of the students' senior year.

Teaching performance was assessed by the college supervisor and cooperating teacher during the fall semester of the four year of students' academic program. Assessments were completed independently, and then shared with the student teacher. Students were informed that the data from their teaching would become a part of the study and they were free to not volunteer this data. The principal investigator assigned the same four digit code to each set of teaching performance assessments. All names of students were eliminated. Then investigators averaged performance scores and entered the data.

Results

Hypothesis One

The third hypothesis of the study is there will be statistically significant positive changes in principled judgment reasoning for the pre-service teacher education sample. Table 3 depicts longitudinal data for principled judgment reasoning for 4-year first-year – senior comparisons with effect size. The table utilizes data organized by McNeel (1994) in his review of college effects. In effect size (d) is used by Pascarella and Terenzini

(1991) in their analysis of college effects: first-year-to-senior change divided by the first-year standard deviation. Bowen (1977) has proposed the following rules of thumb for interpreting effect sizes: small = 0.10 – 0.39, moderate = 0.40-0.69, large = 0.70 – 0.99, and very large = 1.00 and above.

Table 1
Longitudinal Studies of Percent of Principled Reasoning for 4-Year First-Year-Senior Comparisons With Effect Size (d)*

College/University	Longitudinal Results**					
	First Year	Sr	N	SD	Change	d
Universities:						
This study	34.78	47.32	49	21.82	12.54	.57
U Calif., Irvine(A)	36.90	48.10	95	13.26	11.20	.84
West Point (B)	34.40	43.30	104	11.83	8.90	.76
Liberal Arts Colleges:						
Bethel College (C)	35.70	46.40	216	11.62	10.70	.92
Alverno College (D)	35.60	47.40	70	11.53	11.80	1.02
Alverno College (E)	42.80	50.90	70	14.53	8.1	.56
Wheaton College (F)	41.50	52.40	44	17.22	10.90	.63

• In each case, following Pascarella and Terenzini (1991, p.15), effect size is the senior minus first year difference divided by the first-year standard deviation.

** Longitudinal data draws from research by McNeel (1994, p. 32).

A Loxley and Whiteley (1986, pp. 280, 275).

B Bridges and Priest (1983, p. 29).

C McNeel (1994, p.32).

D Mentkowski and Straight (1983, p.178)

E Mentkowski and Straight (1983, p. 178). This sample represents older students.

F Shaver (1987, pp. 214-215).

The results from Table 3 support the third hypothesis. There is longitudinal growth in principled judgment reasoning for the pre-service teacher students. DIT principled reasoning increased from 34.78 to 47.32 over four years. The mean gain for the longitudinal cohort (n=49) was +12.54. The effect size (d) for this change is .57. This is a moderate effect size using the Bowen (1977) rule. Although this is a moderate effect, it is less than the effect sizes for other liberal arts colleges and universities in Table 3. However, some of this difference is accounted for by the large standard deviation for the first-year sample. DIT principled reasoning gain scores for the teacher education sample (n=49, gain = +12.54) was comparable to principled reasoning gains for the other combined longitudinal cohorts (n= , gain= +).

Influence of Gender on DIT Scores

Consistent with other DIT research (Walker, 1991), there were no significant gender differences in principled reasoning. However, there were minimal differences favoring females. Females in the teacher education sample did grow more than the males, however these differences were not statistically significant. In a comprehensive investigation of gender differences on the DIT, Thoma (1986) conducted a meta-analysis of 56 DIT studies involving over 6,000 participants. Thoma found that gender differences accounted for .002 of the variance compared to the variable of education, which was over 250 times more powerful. The outcome of this study contradicts claims made by Gilligan (1977, 1982) that research in this genre is biased against females.

Hypothesis Two

The second hypothesis of the study is that there would be positive correlations between principled judgment reasoning and teaching performance. A Pearson Product-Moment Correlation Coefficient were calculated for participants' DIT (P scores) and Teacher Performance scores. There was a small significant positive correlation ($r=+.15$, $n=49$; $p<.05$). This correlation is somewhat consistent with other findings reported by Rest and Narvaez (1994) who provide statistical summaries of the strength of associations between judgment and actions within a range of 10% to 15%. Although the magnitude of the correlation between DIT –P scores and teaching performance scores was not high in the study reported here, it is quite consistent with other estimates of judgment and action relationships in related fields (Ajzen, 1988). However, some studies show significantly higher correlations. For example, in a study by Duckett and Ryden (1994), first-year DIT scores of entering students predicted an impressive correlation of .58 to clinical performance ratings of nurses in their later years.

As Rest and Narvaez note, their model now embraces four conceptually distinct processes (i.e., Component I: moral sensitivity; Component II: moral judgments, Component III: moral motivation, and Component IV: moral character) which operate together. These other processes may fill the gap between judgments and action, and in doing so, provide a more complete structure for explaining relationships between judgment and action.

Discussion and Implications for Teacher Education

Results of the present study indicate a moderate and significant change in principled judgment reasoning for the pre-service teacher education students. This finding was compared with other four-year longitudinal studies and the results are similar – that is, there were substantial positive gains in moral judgment reasoning during the four years.

As well, correlational results of the present study indicate a low but significant positive correlation ($r=+.15$, $n=49$; $p<.05$) between DIT (P scores) and student teaching performance scores. This finding is consistent with a number of studies that have found a statistical link between scores on measures of principled judgment reasoning and measures of behavior (Rest, Narvaez, Bebeau, & Thoma, 1999). In their review of 47 potential links between the DIT P score and various measures of behavior, they found 32 statistically significant relationships.

Research evidence is mounting to suggest that although teacher education students enter college at lower levels than college students with other majors (Rest & Narvaez, 1994), curriculum can be developed that foster important gains in principled judgment reasoning. In this study, teacher education students participated in additional service learning (role taking) that was joined with intensive self-analysis and reflection. These students showed gains in principled judgment reasoning that were greater than their counterparts in other teacher education programs (McNeel, 1994). McNeel (1994) notes that P-scores for college seniors majoring in business and education have been found to be significantly lower than those of seniors in other majors. Furthermore, seniors in business and education show principled judgment scores that are more like college first-year students (average P-scores in the range of 34 to 42) rather than other college seniors (average P-scores of 43 to 52). The average P score for seniors in the teacher education sample (47.32) were more in keeping with these higher scores.

The promising results of this study lead us to believe that there was a “roletaking and reflection” effect. That is, additional roletaking with guided reflection may account for greater gains in principled judgment reasoning and may fill in some of the gap between judgments and action that has been reported in prior research. The investigators attribute this promising finding to the integration of a roletaking and guided inquiry curriculum intervention during the teacher education students’ four year experience. Examples of complex new roletaking include sustained service learning, tutoring, and other complex human-helping experiences. Applications of this construct and related research to teacher professional development are reviewed by Sprinthall, Reiman, and Thies-Sprinthall, 1996). As well, the “roletaking and reflection” effect has been reviewed by Reiman (1999). while (1999) describes the importance of meta-reflection as a potential fifth component needed for our understanding of moral psychology.

According to Rest, et al. (1999) “The critical characteristic of a college for promoting moral judgment seems to be a commitment to critical reflection” (p.73). Yet, many suggest that teacher education too often fails to integrate awareness of and discussion about ethical issues in teaching (Beyer, 1991; Goodlad 1990, 1994; Hostetler 1997; Oser & Altof, 1993; Sockett, 1993; Strike, 1996). Yet dialogue is not sufficient. Dialogue must be joined with action or as praxis (Friere, 1981). This study suggests that the reflection effect is important, and must be integrated into teacher education programs (as well as other professional preparation programs) in ways that permit future teachers to reason about and respond to the many moral and ethical issues that arise in the context of their progressively more “practice-based” experience (i.e., roletaking).

References

- Ajzen, I. (1988). Attitudes, personality, and behavior. Chicago: The Dorsey Press.
- Beyer, L.E. (1991). Schooling, moral commitment, and the preparation of teachers. Journal of Teacher Education, 42, 205-215.
- Bridges, C., & Priest, R. (1983). Development of value and moral judgments of West Point Cadets (Report No. USMA-OIR-83-002). West Point, NY: United States Military Academy, Office of Institutional Research.

- Chang, F.Y. (1994). School teachers' moral reasoning. In J. Rest and D. Narvaez (Eds.), Moral development in the professions: Psychology and applied ethics (pp. 71-83). Hillsdale, NJ: Erlbaum.
- Davison, M.L. (1979). The internal structure and the psychometric properties of the Defining Issues Test. In J. Rest (Ed.), Development in judging moral issues (pp. 223-245). Minneapolis: University of Minnesota.
- Duckett, L.J., & Ryden, M.B. (1994). Education for ethical nursing practice. In J. Rest and D. Narvaez (Eds.), Moral development in the professions: Psychology and applied ethics (pp. 51-70). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Friere, P. (1981). Education for critical consciousness. New York: Continuum.
- Gilligan, C. (1977). In a different voice: Women's conceptions of self and morality. Harvard Educational Review, 47, 481-517.
- Gilligan, C. (1982). In a different voice: Psychological theory and women's development. Cambridge, MA: Harvard University Press.
- Goodlad, J.I. (1990). Places where teachers are taught. San Francisco: Jossey-Bass.
- Goodlad, J.I. (1994). Educational renewal. San Francisco: Jossey-Bass.
- Hostetler, K. D. (1997). Ethical judgment in teaching. Boston: Allyn and Bacon.
- Katz, L., & Raths, J. (1985). Dispositions as goals for teacher education. Teaching and Teacher Education, 1, 301-307.
- Lampe, J. R. (1994). Teacher education students' moral development and ethical reasoning processes. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.
- Loxley, J. C., & Whiteley, J. M. (1986). Character development in college students. Volume II: The curriculum and longitudinal results. Schenectady, NY: Character Research Press.
- McNeel, S.P. (1994). College teaching and student moral development. In J. Rest & D. Narvaez (Eds.), Moral development in the professions: Psychology and applied ethics (pp. 27-49). NJ: Erlbaum.
- Mentkowski, M., & Straight, M. J. (1983). A longitudinal study of student change in cognitive development, learning styles, and generic abilities in an outcome-centered liberal arts curriculum. Milwaukee, WI: Alverno College, Office of Research and Evaluation. (ERIC Document Reproduction Service, No. ED 239 562)
- Mentkowski, M. and Associates (2000). Learning that lasts. San Francisco: Jossey-Bass.
- Maslovaty, N. (2000). Teachers' choice of teaching strategies. Journal of Moral Education, 29(4), 429-444.
- NCATE (2000, March). NCATE 2000 Unit Standards for Colleges of Education. Washington, D.C.: NCATE
- Oser, F., Dick, A., & Patry, J.-L. (Eds.) (1992). Effective and responsible teaching: A new synthesis. San Francisco: Jossey-Bass.
- Oser, F., & Altof, W. (1993). Trust in advance: On the professional morality of teachers. Journal of Moral Education, 22, 253-272.
- Pascarella, E.T., & Terenzini, P. (1991). How college affects students: Findings and insights from twenty years of research. San Francisco, CA: Jossey-Bass.

Reiman, A.J. (1999). The evolution of the social roletaking and guided reflection framework in teacher education: Recent theory and quantitative synthesis of research. Teaching and Teacher Education, 15(6), 597-612.

Reiman, A.J., Sprinthall, N.A., & Thies-Sprinthall, L. (1999). The conceptual and ethical development of teachers. In R. Mosher, D. Youngman, and J. Day (Eds.), Human development across the life span (pp. 89-102). Westport, CT: Praeger.

Rest, J. (1979). Development in judging moral issues. Minneapolis: University of Minnesota Press.

Rest, J. (1983). Morality. In P.H. Mussen (Series Ed.) & J. Flavell & J. E. Markman (Vol. Eds.), Handbook of child psychology: Vol. 3, Cognitive Development (4th ed., pp. 556-629). New York: Wiley.

Rest, J. & Narvaez, D. (1994). Moral development in the professions: Psychology and applied ethics. Hillsdale, NJ: Lawrence Erlbaum..

Rest, J., Narvaez, D., Bebeau, M., & Thoma, S. (1999). Postconventional moral thinking: A neo-Kohlbergian approach. Mahwah, N.J.: Erlbaum.

Schrader, D. (1999). Metacognitive reflection in university students. In R. Mosher, D. Youngman, and J. Day (Eds.), Human development across the life span (pp. 89-102). Westport, CT: Praeger.

Shaver, D. G. (1985). A longitudinal study of moral development at a conservative religious liberal arts college. Journal of College Student Personnel, 28, 211-218.

Sockett, H. (1993). The moral base for teacher professionalism. New York: Teachers College Press.

Sprinthall, N.A., Reiman, A.J., & Thies-Sprinthall, L. (1996). Teacher professional development. In J. Sikula (Ed.), Second handbook of research on teacher education. New York: Macmillan.

Strike, K. (1996). The moral responsibilities of educators. In J. Sikula (Ed.), Second Handbook of Research on Teacher Education (pp. 869-892). New York: Macmillan.

Thoma, S. (1986). Estimating gender differences in the comprehension and preference of moral issues. Developmental Review, 6, 165-180.

Thoma, S., Rest, J., & Barnett, R. (1986). Moral judgment, behavior, decision making, and attitudes. In J. Rest (Ed.), Moral development: Advances in research and theory (pp. 133-175). New York: Praeger.

Walker, L. (1991). Sex differences in moral reasoning. In W. Kurtines & J. Gewirtz (Eds.), Handbook of moral behavior and development (pp. 333-364). Hillsdale, NJ: Lawrence Erlbaum Associates.

Yost, D.S. (1997). The moral dimensions of teaching and preservice teachers: Can moral dispositions be influenced? Journal of Teacher Education, 48, 281-292.

