

The Accuracy Improvement Measure: A Tool for Assessing the Effectiveness of Teacher Preparation Programs

Mark A. Koorland
University of South Florida-St Petersburg

Sandy MacLeod
University of Utah

Four pre-service teachers trained, and one untrained, in Precision Teaching (PT) were compared employing the Accuracy Improvement Measure (AIM), as an outcome variable. All teachers taught elementary aged students. All five teachers participated in a university practicum. The teachers taught reading, math and related skills for one 3 hour meeting over 10 weeks. Pilot results showed that the AIM was measure suitable for detecting differences among teachers. Description of the AIM, its implications for accountability and examining teacher preparation programs are discussed.

DESCRIPTORS: Teacher preparation, accountability measures, Precision Teaching

The content of teacher training has become more important in the continuing call to reform instruction in the schools (Mathes & Torgesen, 1998; Moats, 1994; Darling-Hammond, 2000). Teacher training is viewed as an important variable that influences what teachers do and how their students learn (Wilson, Floden, & Ferrini-Mundy, 2002). Various methods have been employed to evaluate the effectiveness of teachers including direct observation of teacher practices, teacher responses to questionnaires, and learner change. These measures have been used alone or in combination. A few studies in special education serve to illustrate common quantitative methods of approaching the accountability challenges of teacher effectiveness.

Blackwell, (1972); Meisgeiger, (1965); Sindelar, Espin, Smith, & Harriman, (1990), and Westling, Koorland, & Rose, (1982) all sought to evaluate the relationships among teacher behaviors, teacher preparation, and teacher effectiveness. Various dependent measures were employed. Westling et al. (1981) compared demographic information and instructional practices of exceptional student educators designated as either superior or average. The designation of superior or average was obtained by querying the special educators' district level administrators. Differences between the superior and average teachers on their instructional activities, characteristics, and degree of preparation and training were determined by examining teacher questionnaires. Results indicated superior

teachers differed from their average counterparts in years of training, practicum experience, and certain classroom instructional practices.

The Sindelar et al. (1990) study employed teacher volunteers to examine teacher effectiveness via student achievement measures. Teachers were placed in higher or lower designated groups according to their students' mean achievement gains. Teacher effectiveness was related to specific teacher skill repertoires that also have implications for teacher training. Results, however, did not point to teacher preparation as a factor in effectiveness.

Traditionally, direct measures in the form of celerations have not been employed as a learner outcome measure in published teacher effectiveness studies. Precision teaching (PT) and the use of the Standard Celeration Chart, however, permit measurement of change in learner responding over time. Change in frequencies over a week can be computed and the resulting measure, celeration, is defined as count per minute per week (Pennypacker, Gutierrez., & Lindsley, 2003). Producing change in learner achievement is often considered necessary if we wish to adequately evaluate the effectiveness of teacher preparation programs (Carlson, Hyunshik, & Scholl, 2004).

Pennypacker, et al., (2003) describe the accuracy improvement measure (AIM), a useful measure of change in two celerations simultaneously. The AIM results from describing, as a multiplier, the relationship between concurrent celer

ations of correct and incorrect learner responses. Specifically, incorrect celeration can be used as a reference celeration from which a multiplier is calculated to describe a comparison celeration, in this case, celeration of corrects. The result, a multiplier, is an indicator of the change per week in the accuracy (i.e., the number of correct responses for each incorrect response). Consequently, the greater the value of the AIM produced for a particular learner then the greater the change produced in accuracy per week. Since the AIM is independent of initial accuracy, it permits comparison of change in performance accuracy across learners performing at varied frequencies of correct and incorrect responding (Pennypacker, et al., 2003).

AIMs calculated from concurrent correct and incorrect celerations may be averaged to produce summary measures. For example, all teacher x's charted data can be analyzed and an AIM (based on the learners' correct celerations in relation to incorrect celerations) calculated for each pinpoint the teacher taught. The same calculations can be performed for the charted data of teacher y and a comparison of the average change in accuracy per week produced by each becomes possible. Pennypacker et al. (2003) state that the AIM is useful for comparing the relative effects of teachers, curricula, and programs, and point out that in program evaluations, the AIM has been used to compare learners under various service delivery models in order to determine the model that produces the greatest change relative to the cost of the program.

Our pilot study attempted to employ the AIM for examining the association between learner performance and their pre-service teacher's preparation. Charted data and subsequent AIMs produced by pre-service teachers trained one way were compared to the AIMs produced by a teacher trained differently. Specifically, we noted pre-service training that included PT coursework versus training that did not include PT coursework as a programmatic variable, and examined, for each teacher, their charted data, and associated AIMs as an outcome variable.

METHOD

During practicum training for pre-service teachers enrolled in The Florida State University

Special Education teacher preparation program, an opportunity arose to examine charted data of elementary-aged learners tutored by pre-service practicum students trained in PT (i.e., having taken a 45 contact hour university graduate course in Precision Teaching), and learners taught by one pre-service teacher who did not have the PT course. The non PT trained pre-service teacher received otherwise identical training to the other pre-service teachers. Three teachers taught among five different pupils one semester, and two student teachers taught among another five pupils during a subsequent semester. Each semester, the teachers enrolled in the practicum, taught and timed reading, math, and related skills with their pupils during an evening 3-hour practicum for 10 weeks.

Pre-service Teacher Participants

Pre-service participants were five female graduate students in special education, all with highly similar experience, demographics, and pre-service preparation, except for one teacher, Melody, who had not taken the graduate methods course in PT. The PT course had been taken by the other four teachers. All five pre-service teachers were enrolled in a state approved three year, junior, senior, master's program resulting in exceptional student education licensure. The program required 90 semester credit hours across the three years. In this particular program, both the Master's and Bachelor's degree and were awarded at completion.

The pre-service teachers were enrolled in a graduate level practicum supervised by two practicum supervisors trained in PT. Three of the teachers enrolled in the practicum the first semester of the pilot, and two enrolled during the second semester. Taking the PT course was a prerequisite for enrollment in these particular field experiences. Melody, however, was admitted inadvertently to the second semester PT practicum without having taken the prerequisite PT course. Because of the administrative error, she was allowed to continue to participate. She received assistance from the practicum supervisors on site and fellow students in learning very basic pinpointing, probing, and charting skills as she worked through the practicum.

PT Training

The PT course available to the pre-service teachers addressed the nature of direct versus indirect measures, charting, chart based measures, chart interpretation, and use of the frequency finder. Other topics focused on constructing teacher made probes and using other available materials for obtaining performance samples, setting aims, and using decision rules to prompt instructional changes or adjustments while teaching. Additionally, the course included content about direct instruction. All students in the PT course were required during the 45 contact hours (i.e., 3 semester credit hours) to use PT with a learner (instructional strategies for each learner were determined by the pre-service teacher), write a report, and share charted results with classmates. Students took numerous timed probes in class as one of the course evaluation activities. The charted instructional project comprised the other major course activity.

Learners

Five elementary aged learners, CA 6 to 11, were available each semester (1 male and 4 females during the first semester, 3 males and 2 female during the second semester). The learners had been variously classified with mild learning and behavior disorders. One female during the first semester had undergone psychiatric hospitalization. The learners' caregivers were enrolled in an evening parenting class that met once a week at a state office. That office served as the setting each semester, and contained adequate space for each teacher to work with students individually or in small groups.

Procedures.

The pre-service teachers assessed each child's current levels in reading and math. For the teachers trained in PT, aims were set for tool skills in reading and math. Instructional sessions were conducted interspersed with arts and crafts and games. A point system was employed by all the teachers for behavior management. Learners participated in 1 minute probes each session, and their performance was charted on a Standard Celeration Chart. Table 1 displays the teacher, their students, and each student's pinpoint. Students rotated among teachers, however, each teacher chose the pinpoints for their student, and determined the instructional

activity for the student as well.

Analysis: Celerations and Improvement Measure

For each learner's chart, correct and incorrect frequencies and associated celerations were recorded. Celerations were drawn using the method found in Graf & Lindsley (2002). Each teacher's charts were analyzed in terms of celerations produced for concurrent correct and incorrect responses.

An overall AIM was calculated for each pre-service teacher. Each teacher's charts were analyzed and an AIM was determined for each of their charted pinpoints. Next, all AIMs for a particular teacher were averaged resulting in a mean or overall AIM for each teacher.

RESULTS

Because curricular choice and instructional actions originate with the teacher, then one may generally associate student gains with such choices. Clearly, it is difficult to say with confidence that the overall AIM produced by a teacher is solely the result of his or her training, especially since experience, supervision and other variables could be operating. However, if experiential, supervision and demographic variables are all generally similar across teachers and only their preparation differs, then one could perhaps suggest an approximate association between the teacher training and their instructional efforts in terms of charted progress of students. Without a more carefully designed and managed study than we describe here, only a pilot demonstration of how one might approach studying the relationship between teacher preparation and student outcomes is possible. As a teacher accountability model, the notion of associating summary measures derived AIMs with teacher actions appears reasonable, and due to the properties of celeration and the AIM as an outcome measure, such a model appears feasible.

Comparing the AIMs (Table 2 displays overall AIMs by teacher) between the trained teachers and Melody illustrates this result quantitatively. Melody's charts showed smaller changes in learner accuracy over time than those who were trained differently (i.e., taken the PT course). The overall AIM's in Table 2 were averaged (calculation of geometric means appropriate for ratios

Table 1.
Pre-service Teachers, Student, and Pinpoints

<i>Teacher</i>	<i>Student (CA)</i>	<i>Pinpoint</i>
Melody	Katrina (7)	See/say words in context
	Katrina (7)	See/say words
	James (6)	See/say letter names
	James (6)	See/say phonemic sounds
	Helena (11)	See/say words in context
	Kristen (6)	See/say words
	Richard (8)	See/say words
Rachel	Helena (11)	See/say fractions
	Helena (11)	See/write temperature
	James (6)	Counts backwards
	Katrina (7)	See/say numbers
	Katrina (7)	See/write numbers
	Richard (8)	See/write division
	Kristen (6)	See/say numbers
Andrea	Jason (9)	See/say words (reading passage)
	Zack (7)	See/say words (reading passage)
	Susan (10)	See/say words (reading passage)
	Melissa (9)	See/say words (reading passage)
Brenda	Zack (7)	See/say words
	Susan (10)	See/say words
	Melissa (9)	See/say words
	Jennifer (7)	See/say words
April	Jennifer (7)	See/write one digit numerals
	Zack (7)	See/write multiplication numerals (X2 Table)
	Jason (9)	See/write multiplication numerals (X2 Table)
	Susan (10)	See/write multiplication numerals (X2 Table)
	Melissa (9)	See/write multiplication numerals (X2 Table)

Table 2.
Comparison of Accuracy Improvement Measures (AIM)

<i>Teacher</i>	<i>Mean AIM by Teacher</i>	<i>Mean AIM By Training</i>
Trained in PT		x2.00
Brenda	x2.31	
Rachel	x2.29	
April	x1.98	
Andrea	x1.85	
Untrained in PT		x1.47
Melody	x1.47	

such as the AIM are described in Pennypacker et al., 2003) and produced a mean of $x = 2.0$. Melody's AIM value was $x = 1.47$. Melody's mean AIM is the lowest among the five teachers, and represents a value that is approximately 74 % of the mean AIM ($1.47 / 2.00 \times 100$) for the other four teachers.

DISCUSSION

In this pilot, we were able to note differences among teachers in learner growth in accuracy through examination of their AIMs. Further, differences between the PT trained and untrained pre-service teachers, while not attributable to their training with confidence, are nevertheless interesting to note. Unfortunately, an insufficient number of teachers untrained in PT, do not permit any clear association between training and differential learner growth. What particular component of the PT course or what combination of course components and other experiences that may have contributed to the findings are unknown. Ideally, if teachers could be assigned first to different training models and then their AIM's examined, statements about the function of different teacher preparation strategies could be offered. Since this pilot does not meet the requirements necessary for determining causal relationships, its value may reside in providing a model for approaching teacher preparation program accountability. The method and results suggest a strategy for examining the usefulness of larger units (i.e., a course) of a particular teacher preparation curriculum.

The AIM should be explored further as an accountability measure that may be sensitive to independent variables such as instructor training. Accountability is now regularly sought and expected as an element in programs for learners in public education and other training venues. Additionally, university teacher preparation program accreditation agencies are seeking to obtain data about the effectiveness of teacher preparation programs in terms of outcomes for those who are taught by the programs graduates (Illinois State Board of Education, n.d.). This shift to outcome measures in accreditation presents new challenges to those colleges and schools that in the past needed only to provide curricular, human resource, and program descriptions to the accrediting agency.

What more powerful method to show the

value of training for teachers than to demonstrate the outcomes associated with variously trained teachers? Typically, measures such as certification test passage rates, principal evaluations, and supervisor evaluations are employed to determine the fitness of a beginning teacher. Changes in any measure of learner performance, however, are rare as a tool to determine teacher preparation program efficacy. The AIM provides an easily computed efficacy measure, perhaps, useful in evaluation of a teacher's preparation long before one enters the workforce where other measures, often less direct and sensitive, will prevail.

REFERENCES

- Blackwell, R.B. (1972). Study of effective and ineffective teachers of the trainable mentally retarded. *Exceptional Children, 39*, 139-143.
- Carlson, E., Hyunshik, L., & Scholl, K. (2004). Identifying attributes of high quality special education teachers. *Teacher Education and Special Education, 27*, 4, 350-359.
- Darling-Hammond, L. (2000). Reforming teacher preparation and licensing: Debating the evidence. *Teachers College Record, 102*, 5, 5-27.
- Graf, S. & Lindsley, O. (2002). *Standard celeration charting 2002* (Preliminary ed.). Poland, OH: Graf Implements.
- Illinois State Board of Education. (n.d.). *Certification and professional preparation: NCATE 2000 standards*. Retrieved December 15, 2004, from <http://www.isbe.state.il.us/profprep/ncate.htm>
- Mathes, P. G., & Torgesen, J. K. (1998). All children can learn to read: Critical care for the prevention of reading failure. *Peabody Journal of Education, 73*(3 & 4), 317-340.
- Meisgeiger, C. (1965). The identification of successful teachers of mentally or

physically handicapped children.
Exceptional Children, 32, 229-235.

Moats, L. C. (1994). The missing foundation in teacher education: Knowledge of the structure of spoken and written language. *Annals of Dyslexia, 44*, 81-102.

Pennypacker, H.S., Gutierrez, A., & Lindsley, O. R. (2003). *Handbook of the standard celeration chart*. (Deluxe ed.). H. S. Pennypacker, Gainesville, FL.

Pennypacker, H.S., Koenig, C.H. & Lindsley, O.R. (1972). *Handbook of the standard behavior chart*. H.S. Pennypacker, Gainesville, FL.

Sindelar, P.T., Espin, C.A., Smith, & Harriman, N.E. (1990). A comparison of more and less effective special education teachers in elementary-level programs. *Teacher Education and Special Education, 13 (1)*, 9-16.

Westling, D. L., Koorland, M.A., & Rose, T. L. (1981). Characteristics of superior and average special education teachers, *Exceptional Children, 47(5)*, 357-363.

Wilson, S. M., Floden, F. E., & Ferrini-Mundy, J. (2002). Teacher preparation research: An insider's view from the outside. *Journal of Teacher Education, 53, (3)*. 190-204.