

DO YOU SEE WHAT I SEE? AN EXPLORATION OF SELF-PERCEPTION IN THE CLASSROOM

Eileen Durand Faulkenberry
Texas A&M University - Commerce
Eileen_Faulkenberry@tamu-commerce.edu

Thomas J. Faulkenberry
Texas A&M University - Commerce
Thomas_Faulkenberry@tamu-commerce.edu

Previous research has established a tenuous relationship between what teachers say they do and what teachers actually do in classrooms. In a 3-year longitudinal study, we followed a group of middle-school mathematics teachers in a professional development project. We investigated the relationship between these teachers' self-ratings of standards-based classroom practices and the ratings of a trained observer. We found that over time, the teachers shifted from overestimating the frequency of their standards-based classroom practices to underestimating their standards-based behaviors. The professional development project lessened the gap between what teachers say they do and what they actually do in the classroom.

While the current classroom environment is centered on assessment, there is typically less concentration on teacher self-assessment. Indeed, teacher self-assessment has been found to be somewhat lacking in reliability and validity (Ross, 2006; Burstein et al., 1995; Hook & Rosenshine, 1979). On the other hand, teacher self-assessment has been used positively as a method for promoting teacher efficacy (Ross & Bruce, 2007). Given the equivocal nature of claims regarding teacher self-assessment, especially in relation to assessments from outside observers, we set out to investigate the relationship between teacher self-assessment and actual classroom practices.

Generally speaking, there is a questionable relationship between the classroom practices that teachers report and their actual classroom behaviors. Frykholm (1996) collected 153 classroom observations of 41 preservice mathematics teachers. He found, quite paradoxically, that preservice teachers were readily able to give explanations for why they couldn't use standards-based classroom practices, while simultaneously reporting a high degree of symmetry between the standards and their own teaching. In fact, in a review of research comprising data from over 2300 teachers, Hook and Rosenshine (1979) asserted that "we cannot assume that these [self-] reports correspond to actual practice" (pp. 9-10).

In a more recent study, Ross, McDougall, & Hogaboam-Gray (2003) developed a survey based on several dimensions of standards-based teaching. While they were able to demonstrate that several high scoring teachers were indeed implementing standards-based teaching in their classrooms, they also found that there was still quite a discrepancy between the teachers' perceptions of their standards-based teaching and their actual teaching behaviors. The authors concluded that this discrepancy was due to two main factors: (a) they did not understand what

standards-based teaching meant, or (b) they could not figure out how to implement their ideals into classroom practice. Given that professional development can ameliorate such concerns, it would be interesting to ask whether the same discrepancy between perceptions and actual behaviors exists even after extensive professional development.

The present study was designed to assess the effects of a long-term professional development program on teachers' perceptions of their abilities to teach mathematics using standards-based classroom practices. Critically, we chose to examine this through a comparison of self-reports of standards-based classroom behavior and the reports of a trained classroom observer.

As the previous literature in this area is sparse, predictions are difficult to formulate. However, we did expect to find that, compared to outside observation, teachers tend to overestimate the frequency of their own standards-based classroom behaviors. It is not clear whether this overestimation would persist throughout the duration of a prolonged professional development program.

Method

Participants

Five mathematics teachers from the Northeast Texas region participated in the study. The teachers were participants in a long-term professional development project directed by the principal author. Two of the participants were teachers in a suburban district. The three remaining participants taught in primarily rural districts. The grade levels represented ranged from Grade 4 to Grade 8. The mean years of teaching experience was 9 years ($SD = 6.7$ years, range 2-17 years) at the beginning of the study (2009).

Materials

Participants' standards-based classroom practices were assessed using a classroom observation protocol consisting of 23 items, each representing a statement of a teacher-action or a student-action (Papakonstantinou & Parr, 2004). Example items include "Teacher asks a variety of questions" and "Students are encouraged to explain the process used to reach a solution." Items are scored on a 5-point Likert scale indicating the degree to which a statement occurs during a lesson (1 = low frequency, 5 = high frequency).

Procedure

Over the course of three school years (2009, 2010, 2011), the participants' standards-based classroom practices were assessed in two different ways. First, participants were observed teaching

a lesson in their classrooms by a trained outside observer who completed the classroom observation protocol at the conclusion of the lesson. The same observer rated each participant during each year of the study. Second, each participant then completed the classroom observation protocol as part of a reflective prompt about the lesson they taught. This self-rating was completed within one day of the teaching of the lesson.

Results

For each participant, a mean of 23 classroom observation protocol items was recorded in each of 6 conditions defined by crossing the factors of Year (2009, 2010, 2011) and Rating Type (self-rated, observed). These means were subsequently analyzed using a 2 x 3 within-subjects analysis of variance (see Table 1). Comparisons between means were made using 95% confidence intervals, which were calculated using the method of Loftus and Masson (1994).

The analysis of variance revealed no main effect of Rating Type, $F(1,10) = 0.32, p > 0.6$ and only a marginal main effect of Year, $F(1,10) = 6.47, p = 0.064$. Critically, there was a significant interaction between Rating Type and Year, $F(1,10) = 10.08, p = 0.034$. As can be seen in Figure 1, as the longitudinal trend of self-ratings was stable over the three years, the *observed* ratings steadily increased from 2009 to 2011, reflecting an overall improvement in standards-based teaching practice that was independent of the participants' self-perceptions of their teaching practice.

Table 1.

Mean item scores grouped by Rating Type and Year

Rating type	2009		2010		2011	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Self-rated	3.9	0.6	3.9	0.6	4.0	0.4
Observed	3.3	0.5	3.8	0.5	4.3	0.5

Note. Item scores can range from 1 to 5, where 5 indicates high frequency of standards-based teaching practices.

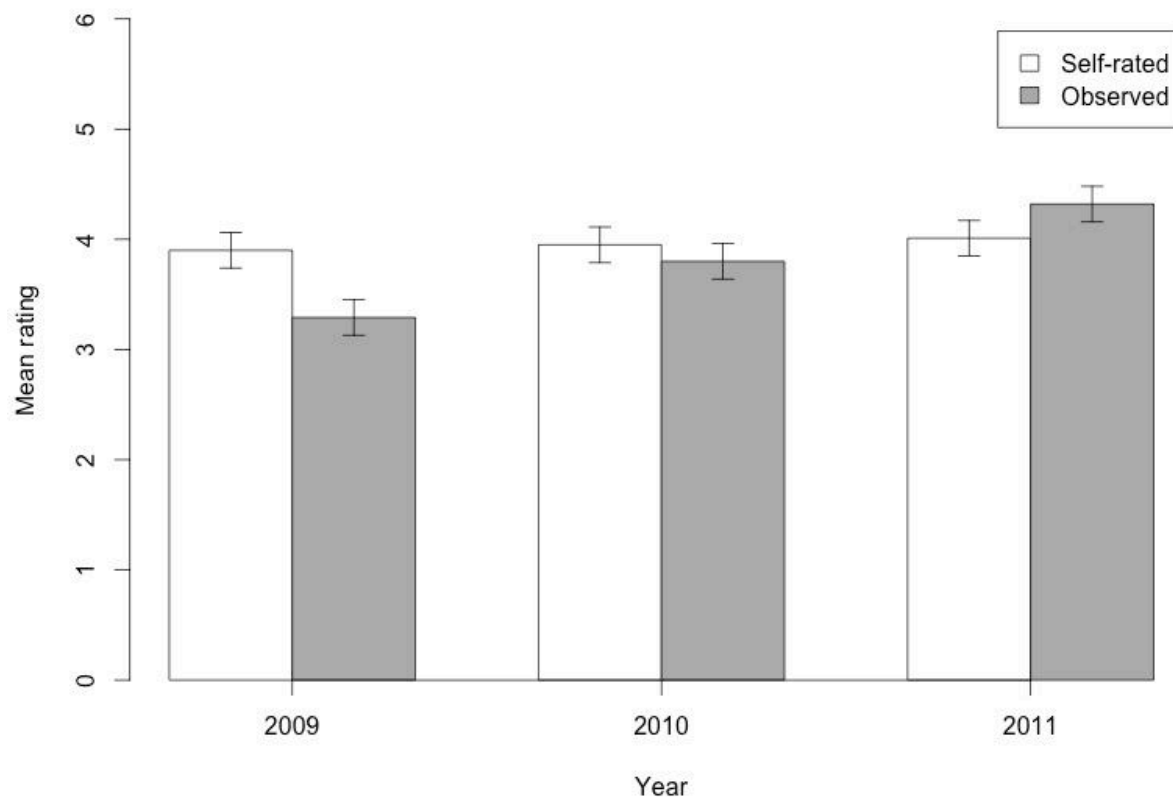


Figure 3: Mean item ratings grouped by Year and Rating Type. The error bars represent 95% confidence intervals, calculated using the method of Loftus & Masson (1994).

It is interesting to note that at the beginning of the long-term professional development project (2009), participants *over-rated* their own standards-based teaching practice a full standard deviation above the ratings of a trained outside observer. However, two years later (2011), the opposite trend prevailed in which participants *under-rated* their own teaching practice almost a full standard deviation *below* the outside rating. That is, as time progressed, participants' abilities to teach using standards-based practices increased beyond their own self-awareness. This trend is remarkable, and to our knowledge, has not been previously found in the literature on teacher development.

Discussion

The present study was designed to assess the effects of a long-term professional development program on teachers' perceptions of their abilities to teach mathematics using standards-based classroom practices. Critically, we chose to examine this through a comparison of self-reports of standards-based classroom behavior and the reports of a trained classroom observer. Predictions

based on previous literature were difficult to formulate, but we did predict a general overestimation of teachers' perceptions of their own standards-based classroom practices. Perhaps surprisingly, we did not find this. Compared to the ratings of a trained outside observer, there was no overall difference between teachers' self-rated classroom practices and the observed classroom practices.

When adding the additional variable of year, however, the data becomes more interesting. Indeed, we found a significant interaction between year and rating type. That is, as the time in the professional development program increased, teachers moved from overestimating the frequency of their standards-based classroom practices (relative to a trained outside observer) to underestimating their standards-based practices. This pattern is remarkable and was not expected.

Specifically, this interaction was due to differences in the longitudinal trends between the self-ratings and the observed ratings. Whereas self-rated standards-based classroom behaviors remained constant over the three years of the professional development program, the actual observed behaviors increased steadily. Although this growth in observed standards-based teacher practices is a testament to the efficacy of the professional development program, it is still puzzling as to why the self-ratings were stable. Two possible explanations are given below.

It is conceivable that the professional development program, by targeting specific teaching behaviors related to standards-based classroom practice, is truly effective in promoting gains in teacher efficacy while leaving teacher self-beliefs unchanged. However, data exists for this specific professional development program that calls this explanation into question. Specifically, we have seen modest, reliable gains in teacher self-efficacy over the course of the professional development program. These data are not tied to the current study, so further conclusions on this matter are left unwarranted until further studies can be conducted.

Another possibility is that over the course of professional development, teacher self-ratings are subject to a suppressive effect whereby teachers reliably underestimate their own efficacy as they become better teachers. The current data support this view but do not rule out alternative explanations, such as maturation. The case for a maturation-only account is weakened somewhat by the range of years of experience in our study (2-17 years). But nonetheless, future studies will need to carefully consider a manipulation of the professional development program, possibly by following a group of teachers who were not in such a program.

In conclusion, we examined the effects of a long-term professional development program on teachers' perceptions of their abilities to teach mathematics using standards-based classroom

practices. Through a comparison of self-reports of standards-based classroom behavior and the reports of a trained classroom observer, we found that over time, teachers shifted from a general tendency to overestimate their abilities to an underestimating tendency. The exact reason for this shift is unclear, but it is primarily due to the teachers' increased standards-based classroom behaviors through a long-term program of professional development.

References

- Burstein, G. W., McDonnell, L. M., van Winkle, J., Ormseth, T., Mirocha, J., & Guitton, G. (1995). *Validating national curriculum indicators*. Santa Monica, CA: RAND.
- Frykholm, J. A. (1996). Preservice teachers and mathematics: Struggling with the Standards. *Teaching and Teacher Education, 12*, 665-681.
- Hook, C. M., & Rosenshine, B. V. (1979). Accuracy of teacher reports of their classroom behavior. *Review of Educational Research, 49*, 1-12.
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subjects designs. *Psychonomic Bulletin & Review, 1*, 476-490.
- Papakonstantinou, A., & Parr, R. (2004). Rice University school mathematics project: Geometry module. Retrieved from <http://math.rice.edu/~rusmp/geometrymodule/index.htm>.
- Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, & Evaluation, 11*, 1-13.
- Ross, J. A., & Bruce, C. D. (2007). Teacher self-assessment: A mechanism for facilitating professional growth. *Teaching and Teacher Education, 23*, 146-159.
- Ross, J. A., McDougall, D., & Hogaboam-Gray, A. (2003). A survey measuring elementary-teachers' implementation of standards-based mathematics teaching. *Journal for Research in Mathematics Education, 34*, 344-363.