Parent Surveys for Teacher Evaluation

Kenneth D. Peterson · Christine Wahlquist · Julie Esparza Brown · Swapna Mukhopadhyay

Abstract Parent or guardian perceptions play a specialized role in the evaluation of school teachers. Parents are important stakeholders in teacher success, they are in some instances partners in the teachers’ work, parents have unique personal information about student learning, and they can report on the teacher duties to inform parents about the classroom and child progress. This study analyzed the responses of parents to 12 survey items concerning teacher performance in 201 classrooms. The surveys were used as part of an innovative teacher evaluation program in which teachers elected to include parent feedback as one objective data source for annual review. In this study three factors emerged as important concerns for parents: humane treatment of students, support for pupil learning, and effective communication and collaboration with parents. Recommendations for use of specific survey items can be based on the empirical results of this sampling. The data gathered by parent surveys define one dimension of quality which may vary in importance from one teacher to another.

Keywords

DOI 10.1007/s11092-006-5740-9

K. D. Peterson (✉)
School of Education,
Portland State University, P.O. Box 751, Portland, OR 97207-0751, USA
e-mail: petersk@pdx.edu

C. Wahlquist
Davis School District
45 E. State Street, Farmington, UT 84025, USA

J. E. Brown · S. Mukhopadhyay
School of Education,
Portland State University,
P.O. Box 751, Portland, OR 97207-0751, USA
e-mail: jebrown@pdx.edu; swapna@pdx.edu
Educators describe four ways in which teachers’ work with parents and guardians of their students can be related to teacher evaluation. One teacher duty is to inform parents about the education and progress of their individual children (Dwyer, 1995; Scriven, 1988). Second, teachers are partners and collaborators with parents in pupil learning (Drake & Roe, 1994). A third connection is that parents have personal knowledge about their child’s education which can furnish information about the quality of teacher performance (Epstein, 1985; Peterson, 2000; Stronge & Ostrander, 1997). Finally, parents constitute a distinct stakeholder group which should be included in school system evaluation (Mark & Shotland, 1985).

Because parents are part of the work of teachers, and see it from an important and unique point of view, there is a need to somehow incorporate parent perceptions into teacher evaluation systems in order to get the broadest look at teacher quality. While parents potentially are an important consideration in teacher evaluation, teacher evaluation systems have not often systematically incorporated parent views (Loup et al., 1996; Peterson, 2004). However, recent developments highlight the potential of parent data (Ostrander, 1995; Peterson, 1989a,b; Stronge & Ostrander, 1997) and demonstrate the need for continued research and development for this data source.

The purpose of this study was to refine a survey procedure and items to involve parental views in the evaluation of public school teachers. The development processes were to: a) field test a survey instrument; b) empirically examine survey items; c) determine norms to interpret future use of parent views; d) field test a cut-score strategy; and e) assess teacher satisfaction, with the strategies.

**Background**

Teacher evaluation has not had a strong history of systematically including parent views. For example, McGreal (1983) discouraged the effort, saying that “it would seem that the potential outcomes from parent involvement in teacher evaluation are not significant enough to outweigh the logistical and political implications” (p. 126). Some authoritative guides to teacher evaluation practice (e.g., Millman & Darling-Hammond, 1990; Shinkfield & Stufflebeam, 1995) did not specify methods for including parents. While principals are expected to consider parents as a part of teacher evaluation, very few school districts use systematic parent input in their teacher assessment systems (Lawton, 1997). Parent surveys were not included in the eight methods of teacher evaluation reported in a study of the 100 largest school districts in the United States (Loup et al., 1996). Two authoritative texts on teacher evaluation do include a chapter on client surveys (Peterson, 2000; Stronge, 1997).

The most common current practice in school district teacher evaluation systems is to include parent relations as a checklist item or category in a principal’s evaluation report. Items like “Shows contact [with parents] about positive performance” and “Shows contact [with parents] about possible...”
Problems” (Jordan School District, 1995) encompass recognized duties or expectations for teachers. However, principal reports of teacher performance have been criticized as less accurate and more subject to administrator bias and halo effect ratings than more direct information like parent surveys or focus groups (Stronge & Ostrander, 1997).

Conceptual Basis for Including Parent Views in Teacher Evaluation

While using systematically gathered parent views in teacher evaluation is an uncommon practice, many writers have argued for including them. Scriven (1988) described reporting student progress to parents as a systematic job responsibility of teachers. Dwyer (1995) discussed the expectation of communicating with parents about student learning as contained in the Educational Testing Service Praxis III Professional Assessments. Peterson (1984, 1987, 2000) added to these expectations the opportunity for teachers to be partners with parents by giving them ideas for home support of learning, e.g., reading with them, creating places to do homework, or simply listening to them. He also recommended the use of a global satisfaction rating item (“overall satisfaction with this teacher”) to be included because of the stakeholder status of parents.

Technical Issues Concerning Parent Views in Teacher Evaluation

Epstein (1985) concluded that parent views can be an important source of information about teachers. She reported a correlation of $r = 0.27$ between elementary school principal ratings and those of parents. She found that parent ratings are influenced by student reports of classroom life and by resources and ideas given to the parents by the teacher. Epstein reported that parents of older students rated teachers more positively than did parents of younger children.

Peterson (1987) advocated parent views as an important but distinct data source for teacher evaluation. He found correlations between parent ratings and other data sources as follows: $r = 0.09$ with administrator ratings, $0.58$ with pupil surveys, $-0.13$ with teacher test scores, $-0.13$ with years of teacher experience, and $0.04$ with number of teacher professional activities. He concluded that “administrator reports, the conventional practice, showed low correlations with other measures” (p. 316) of teacher quality.

Peterson (1989a,b) found statistically significant difference in return rates and ratings of parent surveys according to the age of the pupils: parents of younger students returned surveys at higher rates and were more positive about teacher performance. He found that some teachers had stable patterns of parent ratings over several years, while other teachers varied according to the particular year. Use of parent reports was not associated with an improvement in ratings in subsequent years. Peterson presented data about costs, sampling, and use of a global rating item in parent surveys.

Ostrander (1995) found that parent ratings of teachers did not agree with teacher self-ratings and administrator ratings. She reported that parents...
viewed teachers in a positive light, but rated teachers lower than did the administrators. She found a significant degree of agreement between students’ and parents’ rank ordering of teachers. Ostrander concluded that “...the fairest and most comprehensive performance appraisals may involve multiple judges, each offering a unique perspective on teacher effectiveness” (p. ii).

Stronge and Ostrander (1997) recommended client surveys, including parents, as a state-of-the-art remedy for deficiencies in most current teacher evaluation practices. The authors argued for collaboration, legitimacy of judgment, and practicality of data gathering; they presented several prototypical forms.

The logistics of including parent views in teacher evaluation are considerable (McGreal, 1983). Time and dollar costs for complex data collection require deliberation. Peterson (1989a,b) reported that the expense of parent surveys averaged out over all teachers in a district to $15. per teacher. Time costs to teachers are crucial; and time, perhaps, is a teacher’s most valuable commodity (Lortie, 1975). Data gathering requires time and standardized procedures. Forms must be created, duplicated, distributed, and recovered. Issues of privacy, accountability, coding, security, and scheduling require planning and expense. Scoring of parent surveys involves choices of hand or machine approaches. Also, if surveys are scored by hand, it is important that the forms not be tabulated by the teachers themselves, but rather by clerks who must themselves be managed and monitored. Lastly, record keeping must balance the competing concerns of confidentiality, credibility, and costs.

Sociological and Political Issues

While the above studies address the technical questions of parent surveys, teacher evaluation also has sociological and political issues. The use of parent views in teacher evaluation is controversial both sociologically and politically (Epstein, 1985; House, 1973; Lawton, 1997; Lortie, 1975; Peterson, 1984, 2000; Waller, 1932). Teachers disagree about the importance of parent opinions of their work (Kauchak et al., 1985). Some teachers are in a better position to receive favorable parent reactions (e.g., primary grade teachers), while other teachers work quite well out of the view of parents (high school level). Some good teachers include parents as part of their good practice (e.g., as volunteer aides), while other good teachers work quite independently of parents. Some teachers derive power from parent advocates, while other teachers face threats from the same audience. Some teachers with high status in evaluation systems that do not include parents may see the introduction of parent views as a threat to their current status. Finally, some parents who want a role in school-based decision making constitute a powerful external political force in the educational system. Each school differs in how educators foster constructive dialogue and input with their parents.

Needs for Study

Parents are a potential source of important information for teacher evaluation. There is a considerable need for study of procedures, survey items, and policies about increased use in school systems.
The Study

Questions Addressed in the Study

This study addressed development of valid, reliable, and practical strategies for including parent surveys in teacher evaluation. These considerations included selection of survey items, use of a single “global item”, methods of reporting scores to teachers and administrators, selection of logistical practices, estimates of instrument reliability, and sampling of teacher satisfaction with procedures.

Item selection is an important issue for survey development. While evaluators want to request information from parents on many different topics, not all survey items put before them work equally well. Peterson (2000) described a variety of criteria for judging whether or not possible items are acceptable or valid for use with parents. These tests include logical integrity, reasonableness, fairness, and empirical trial. It is important that parents be asked to report conditions of learning that they can directly observe, for example, that the level of work is appropriately challenging (not overwhelming or boring) for their child. It is equally important that parents not be asked to report on conditions they cannot see, for example that the teacher uses time well in class. Fairness of items is important because teachers work in varying performance conditions. For example, some teachers deal with 25 children per day while others face 125; it could be argued that it is unfair to have parents rate the availability of teachers who have different demands placed upon them. This study empirically tested a variety of items for use in the survey to be adopted by a school district.

A second issue studied was the strategy of using a “global item” (Peterson et al., 1984) to make summative judgments about parent views. It is possible to present parent views by reporting each individual item. However, not all item topics are equal in value. For example, the item “presents ideas for home support of learning” may or may not be as important as “creates a good challenge for students, not too hard or too easy.” Several ways to present survey results are: total scale score, average item score, factor score, and weighted item score. The problem with these alternatives is that each presents an inherent value in balancing the contribution of individual items which may distort the issue of quality. A single global item, such as “this is a good teacher” or “rate overall performance,” can be a useful summative report, if the item can be shown to well represent other items, underlying factors, internal reliability, and various possible scales.

The method of reporting parent survey scores is a crucial question for their successful use. Summary statistics (mean, standard deviation, number of responses) on each item give absolute information about parents’ responses. However, the perspective of comparison with other teachers is needed for summative evaluation. Thus, District norms for item response and percent return rate are needed for each grade level. A related problem of reporting parent survey results to teachers is that teachers have difficulty accepting scores that are not perfect (Lortie, 1975; Peterson, 2000). In this present study,
a categorical report ("well functioning" or "not well functioning") was used to discriminate among teachers, but avoid overly narrow comparisons enabled by precise scale scores. A third reporting problem is that the use of categories calls for a cut off score. Possibilities for a cut off score are: some absolute number, a score that defines an absolute percent, a scree test that eliminates laggards, or a population parameter such as a standard deviation below mean. This study examined the latter strategy for satisfactory discrimination in reporting.

The issue of validity and reliability for the teacher evaluation system is not only research related but reflects state law (Utah [Utah State Code, 1953]) where the school district in this study was located. The argument that the larger evaluation system is valid and reliable is partly based on the argument that each of the component surveys and reports is valid and reliable. This study examined whether parent views are a valid and reliable component of the larger system. While school districts are generally not held to standards of perfection in their teacher evaluation systems (Chance v. Board of Examiners, 1971; McCoy, 1998), it is a reasonable expectation that they strive for state-of-the-art techniques, test their assertions of validity and reliability, and document the results of their actual practice.

The Davis School District Educator Assessment System (EAS)

The Davis School District Educator Assessment System (EAS) (Davis School District, 2005) design provides multiple and variable data sources for teacher evaluation. Each teacher must present a constellation of information that their performance is "well functioning." The varied configuration of data types allows for documentation of good teaching of different types and in different settings. It also addresses the political (McGreal, 1983) and sociological (Lortie, 1975) problem of teacher acceptance of data sources. Detractors of any individual data source (e.g., in the case of parent surveys: "parents don’t take the needs of the class into account, just the individual child") are not forced to take a stand which conflicts with the more numerous teacher supporters of parent data. This provision for teacher choice avoids the problem of "all or nothing" data source selection wherein most school districts gather parent views for none of the teachers. An additional feature of the EAS system is that teachers get two levels of control: teachers first must elect to have the surveys collected, and second they must elect to present the results after inspecting them.

The EAS requires that each year teachers furnish data about their performance to their building administrators, who then complete teacher evaluation reports. In addition, a more extensive review that includes at least three different data sources selected from the menu of options is required for: (a) veteran teachers every four years; (b) beginning teachers each year; and (c) teachers on "performance assistance" remedial status. Teachers are assisted in their data collection by District EAS staff, i.e., information like that provided in parent surveys takes teachers no more than five or ten minutes to consider, schedule, and inspect results. Teachers select from a menu of possible data sources that, in their judgment, best represents their value, merit, impact,
and quality. Data sources in addition to parent surveys include student surveys, teacher tests, pupil achievement data, documentation of professional activity conceptually linked to performance, peer review of portfolios, and information unique to an individual teacher. No claims are made that any single data sources work well for, or should be required for, every teacher in the District. The purposes of this complex system are to: (a) have teachers more involved in their own evaluation (Peterson & Chenoweth, 1992); (b) base teacher evaluation on the best objective evidence available; (c) gather information about performance to use in public relations, staff development, and dissemination of best practices; (d) replace time consuming—but not respected (Kauchak et al., 1985; Wolf, 1973)—annual formal classroom evaluation visits by principal; and (e) include legitimate stakeholders in the teacher evaluation system (Mark & Shotland, 1985).

This study was designed to investigate the use and results of one optional data source of the EAS: parent surveys. The goal was to test specific items, overall scoring procedures, and teacher satisfaction with the multiple and variable data evaluation system.

Methods

This study used surveys with 12 scale items to assess the views of parents and guardians of public school students concerning the performance of their teachers. Each item was presented with a three-point scale: 3 = yes, 2 = somewhat, 1 = no. Item descriptors and descriptive statistics are presented in Table 1.

Sample

Parent surveys were elected as a data source for evaluation by 341 teachers from 27 schools in this school district pilot study. Item and grade-level analyses was performed on a total of 3,487 useable surveys. Class and teacher specific

<table>
<thead>
<tr>
<th>Item (descriptor)</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Kurtosis</th>
<th>Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Child is learning</td>
<td>2.87</td>
<td>0.36</td>
<td>7.05</td>
<td>-2.74</td>
</tr>
<tr>
<td>2. Child knows what is expected</td>
<td>2.86</td>
<td>0.37</td>
<td>5.68</td>
<td>-2.52</td>
</tr>
<tr>
<td>3. Work is right challenge</td>
<td>2.74</td>
<td>0.51</td>
<td>2.40</td>
<td>-1.80</td>
</tr>
<tr>
<td>4. Overall satisfaction</td>
<td>2.76</td>
<td>0.52</td>
<td>3.61</td>
<td>-2.13</td>
</tr>
<tr>
<td>5. Child treated as individual</td>
<td>2.85</td>
<td>0.42</td>
<td>7.85</td>
<td>-2.88</td>
</tr>
<tr>
<td>6. Parent aware of expectations</td>
<td>2.68</td>
<td>0.58</td>
<td>1.61</td>
<td>-1.64</td>
</tr>
<tr>
<td>7. Home learning support ideas</td>
<td>2.49</td>
<td>0.71</td>
<td>-0.33</td>
<td>-1.03</td>
</tr>
<tr>
<td>8. Feedback on child progress</td>
<td>2.67</td>
<td>0.60</td>
<td>1.49</td>
<td>-1.63</td>
</tr>
<tr>
<td>9. Caring/respect for child</td>
<td>2.84</td>
<td>0.44</td>
<td>7.01</td>
<td>-2.74</td>
</tr>
<tr>
<td>10. Activities relevant to progress</td>
<td>2.84</td>
<td>0.41</td>
<td>6.24</td>
<td>-2.58</td>
</tr>
<tr>
<td>11. Accessible, responds timely</td>
<td>2.86</td>
<td>0.40</td>
<td>8.70</td>
<td>-3.00</td>
</tr>
<tr>
<td>12. Understands child needs</td>
<td>2.75</td>
<td>0.52</td>
<td>3.09</td>
<td>-1.99</td>
</tr>
</tbody>
</table>
analysis were performed for the 201 teachers who had at least 12 surveys returned per class.

Survey Data and Analysis

Data for this study consisted of independent variables of teacher, level taught, and school. The dependent variables were scores on a 3-point Likert scale for 12 items. Data were collected on bubble mark, computer-sensed forms. Descriptive statistics were computed for the total sample of survey items (mean, standard deviation, kurtosis, and skewness). Multivariate analyses of the 12 items consisted of an intercorrelation matrix and a principal components analysis with a varimax rotation. The internal consistency reliability of the 12 item scale was computed using the Cronbach’s Alpha procedure. Individual analysis of the global item was done with an analysis of variance to assess the effects of school and level of teaching.

Teacher and Principal Satisfaction Data

Participant satisfaction was assessed with a survey of a random sample of 561 teachers, equally representing elementary, middle and high school levels, were surveyed about their satisfaction with the pilot teacher evaluation system, which included the opportunity to sample parent views. Administrator perspectives were documented with interviews and focus groups.

Findings

Descriptive Statistics of Survey Items

Table 1 presents descriptive statistics for each survey item. Means ranged from 2.87 (“child is learning”) to 2.49 (“teacher provides ideas for home support of learning”) on the three point scale (with “1” labeled “No,” “2” labeled “Somewhat,” and “3” labeled “Yes”). All distributions were lepto-kurtic, except for “home support of learning*,” and skewed toward high levels of satisfaction.

Multivariate Analysis of Items

Table 2 shows the intercorrelations of survey items. Individual item correlations ranged from $r = 0.26$ (“Child knows what is expected” and “Teacher provides ideas for home support of learning”) to $r = 0.69$ (“Child treated as individual” and “Caring/respect shown for child”). All correlations were positive, and 57.5% were at, or exceeded, 0.40.

The principal components analysis of the 12 items showed three underlying factors (eigenvalues over 1.0) in the parental responses. These factors are presented in Table 3. Loadings (Pearson product-moment correlations) of each item on the three factors are presented in Table 4. Factor I (eigenvalue 5.86,
48.8% of variance) was Humane treatment of child (relatively highly correlated with “Demonstrates caring and respect for my son/daughter,” “Treats son/daughter as an individual,” “Understands son/daughter’s needs”). Factor II (eigenvalue 1.10, 9.2% of variance) was Supports pupil achievement (“Classroom work is the right challenge,” “Learning activities are relevant to

Table 2  Correlation matrix of survey items

<table>
<thead>
<tr>
<th>ITEM</th>
<th>I1</th>
<th>I2</th>
<th>I3</th>
<th>I4</th>
<th>I5</th>
<th>I6</th>
<th>I7</th>
<th>I8</th>
<th>I9</th>
<th>I10</th>
<th>I11</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITEM2</td>
<td>0.39</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITEM3</td>
<td>0.45</td>
<td>0.34</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITEM4</td>
<td>0.58</td>
<td>0.48</td>
<td>0.46</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITEM5</td>
<td>0.40</td>
<td>0.32</td>
<td>0.28</td>
<td>0.54</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITEM6</td>
<td>0.39</td>
<td>0.40</td>
<td>0.35</td>
<td>0.47</td>
<td>0.41</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITEM7</td>
<td>0.35</td>
<td>0.26</td>
<td>0.35</td>
<td>0.44</td>
<td>0.36</td>
<td>0.59</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITEM8</td>
<td>0.38</td>
<td>0.29</td>
<td>0.32</td>
<td>0.45</td>
<td>0.44</td>
<td>0.59</td>
<td>0.59</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITEM9</td>
<td>0.41</td>
<td>0.35</td>
<td>0.30</td>
<td>0.60</td>
<td>0.69</td>
<td>0.41</td>
<td>0.39</td>
<td>0.44</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITEM10</td>
<td>0.49</td>
<td>0.35</td>
<td>0.53</td>
<td>0.52</td>
<td>0.35</td>
<td>0.37</td>
<td>0.35</td>
<td>0.34</td>
<td>0.39</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>ITEM11</td>
<td>0.35</td>
<td>0.30</td>
<td>0.30</td>
<td>0.48</td>
<td>0.47</td>
<td>0.45</td>
<td>0.41</td>
<td>0.45</td>
<td>0.56</td>
<td>0.38</td>
<td>-</td>
</tr>
<tr>
<td>ITEM12</td>
<td>0.47</td>
<td>0.45</td>
<td>0.39</td>
<td>0.67</td>
<td>0.61</td>
<td>0.50</td>
<td>0.47</td>
<td>0.50</td>
<td>0.68</td>
<td>0.48</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Table 3  Item analysis for parent surveys

<table>
<thead>
<tr>
<th>Item</th>
<th>Load factor1</th>
<th>Load factor 2</th>
<th>Load factor 3</th>
<th>Corr with total</th>
<th>Corr with item 4</th>
<th>Alpha if deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Child is learning</td>
<td>0.30</td>
<td>0.70</td>
<td>0.17</td>
<td>0.58</td>
<td>-</td>
<td>0.58</td>
</tr>
<tr>
<td>2. Child knows what expected</td>
<td>0.28</td>
<td>0.57</td>
<td>0.14</td>
<td>0.52</td>
<td>-</td>
<td>0.48</td>
</tr>
<tr>
<td>3. Work is right challenge</td>
<td>0.03</td>
<td>0.78</td>
<td>0.23</td>
<td>0.52</td>
<td>-</td>
<td>0.46</td>
</tr>
<tr>
<td>4. Overall satisfaction</td>
<td>0.36</td>
<td>0.58</td>
<td>0.23</td>
<td>0.74</td>
<td>-</td>
<td>0.54</td>
</tr>
<tr>
<td>5. Child treated as individual</td>
<td>0.81</td>
<td>0.18</td>
<td>0.19</td>
<td>0.62</td>
<td>-</td>
<td>0.54</td>
</tr>
<tr>
<td>6. Parent aware of expectations</td>
<td>0.25</td>
<td>0.28</td>
<td>0.75</td>
<td>0.66</td>
<td>-</td>
<td>0.47</td>
</tr>
<tr>
<td>7. Home learning support ideas</td>
<td>0.38</td>
<td>0.21</td>
<td>0.82</td>
<td>0.59</td>
<td>-</td>
<td>0.44</td>
</tr>
<tr>
<td>8. Feedback on child progress</td>
<td>0.31</td>
<td>0.18</td>
<td>0.77</td>
<td>0.63</td>
<td>-</td>
<td>0.46</td>
</tr>
<tr>
<td>9. Caring/respect for child</td>
<td>0.85</td>
<td>0.22</td>
<td>0.19</td>
<td>0.65</td>
<td>-</td>
<td>0.59</td>
</tr>
<tr>
<td>10. Activities relevant to progress</td>
<td>0.22</td>
<td>0.74</td>
<td>0.18</td>
<td>0.59</td>
<td>-</td>
<td>0.52</td>
</tr>
<tr>
<td>11. Accessible, responds timely</td>
<td>0.61</td>
<td>0.18</td>
<td>0.37</td>
<td>0.59</td>
<td>-</td>
<td>0.48</td>
</tr>
<tr>
<td>12. Understands child needs</td>
<td>0.70</td>
<td>0.38</td>
<td>0.30</td>
<td>0.73</td>
<td>-</td>
<td>0.66</td>
</tr>
</tbody>
</table>
progress,” “Son/daughter is learning in this class”). Factor III (eigenvalue 1.00, 8.4% of variance) was Communicates/collaborates with parents (“Gives ideas for home support of learning,” “Gives feedback on son/daughter progress,” “Parents are aware of class and pupil expectations”).

The internal consistency reliability analysis showed a Cronbach’s Alpha of 0.895. The effect on Alpha of each item is shown in Table 3. The item having the greatest positive effect on Alpha was the global item while the item having the least positive effect was “Work is right challenge.”

Global Item

Item 4 (“overall satisfaction,” considered a global item), relative to other 11 items, showed the: (a) highest average loading (correlation) on the three factors (.46), (b) highest correlation with total of 12 items (0.74), (c) highest average correlation with each of other 11 items (0.52), and (d) greatest affect on alpha if deleted (0.879).

A multiple classification analysis was performed to estimate contributions of variance by school, grade level, and teacher. The independent variable of

Table 4 Teacher satisfaction with new data system (which includes parents)

<table>
<thead>
<tr>
<th>Item</th>
<th>Agree</th>
<th>Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I had more control over my evaluation</td>
<td>463 (82.5%)</td>
<td>98 (17.5%)</td>
</tr>
<tr>
<td>Better helped me reflect on my teaching</td>
<td>471 (82.5%)</td>
<td>100 (17.5%)</td>
</tr>
<tr>
<td>Improvement over old data system</td>
<td>476 (84.5%)</td>
<td>87 (15.5%)</td>
</tr>
<tr>
<td>Yearly formal observations are not necessary for experienced, well functioning teachers</td>
<td>482 (83.2%)</td>
<td>97 (16.8%)</td>
</tr>
</tbody>
</table>
teacher explained 17.86% (eta squared) of the variance in the dependent variable of the global item.

Cut Scores

The EAS reports out one of two categories to teachers on their global item mean: “Well functioning” or “Not well functioning.” As depicted in Fig. 1, a cut score of 1 standard deviation below the mean excluded 6.0% of the teachers from the “Well Functioning” category report. A cut score of 1.5 standard deviations below the mean would exclude 3.0% of this sample of teachers.

Costs

Costs for the parent survey included computer scored forms and clerical costs for form preparation, distribution, and processing. The forms themselves cost approximately $0.05/form, or $1.50/class. About twice as many forms were prepared as were ultimately reported. The approximate total cost was $12.50 per class.

Participant Satisfaction

Table 4 reports four items from the teacher satisfaction survey. In general, teachers were satisfied with the opportunity to gather data of their choosing, including parent views. This finding is important in light of the strong sentiment that formal administrator observations are not necessary for experienced, well functioning teachers.

Focus group analysis suggested that administrators were equally positive about the new evaluation system, including the use of parent surveys. However, principals were more likely to point out potential problems, such as inadequates sampling and parents not taking the process seriously. They also reported that they learned little from the surveys that they did not already know.

Discussion, Conclusions, and Recommendations

This study supported the contention that parent surveys can be defended as valid and reliable data sources for teacher evaluation. The patterns of response, especially disclosed by the factor analysis, suggest that parents responded to the range of items with reason, intent, and consistent values. Thus, data gathered by parent surveys define one dimension of teacher quality as expressed by parent views, which may vary in its importance from one teacher to another. The positive response of teachers and administrators found in this study is important in light of the general dissatisfaction that teachers have with evaluation practices solely based on administrator reports (Kauchak et al., 1985; Lortie, 1975; Stronge & Ostrander, 1997; Wolf, 1973).
The limits of parent responses should be taken into account in making final judgments based upon them. For example, the amount of variance explained by teacher was high (relative to dependent variables in educational research), but certainly not absolute: error terms and unknown sources of variance remain substantial: For this reason, and others, it is difficult to explain the reasons for low scores. Cautions should be used: high parent ratings do not necessarily mean the same thing as good teaching. Perhaps the best interpretation is that high parent ratings in conjunction with at least several other positive indicators is a good indicator of quality teaching. This view is consistent with that expressed by Glass (1974), Peterson (1984), Epstein (1985), and Ostrander (1995) that multiple judges (or, multiple data sources) are required to best identify teacher quality.

The representativeness of the global item strategy was supported by the findings of this study. That is, evidence was found that the most central views of the parents could be compressed by recognizing the responses to the most conceptually pertinent item (“Overall satisfaction with this teacher”) rather than a scale composite (e.g., gross total of items, weighted scale, factor score) or another single item. Each scale composite suffers in comparison to the global item because of problems such as (a) counting all items as equally important (e.g., “Causes learning” vs. “Is immediately available when called upon”), (b) claims that one weighting scheme (yours) is always better than another (mine), or (c) one underlying scale factor (kind, humane treatment) is always more important than another (gives clear instruction).

Not all of the items used in this pilot study of parent surveys are recommended for use in teacher evaluation. For example, “Child is learning” is conceptually important, but parents are hardly in the position to be adequate reporters of actual teacher performance in this regard. Parents are (a) not in the classroom, (b) ignorant of important comparative data among students, (c) experts about their child but not in classroom teaching and learning, (d) demonstrably swayed by the emotional reaction and success of their individual child rather than the overall success of the classroom, and (e) overly influenced by their own school history (Lortie, 1975). A list of recommended items appears as Appendix A.

One surprising set of findings led to a change in recommendations for specific items to include in a parent survey. The original survey designed for this study did not include “popularity” items (e.g., “Child treated as an individual”) which may lead to ratings more based upon trivial and superficial pleasing teacher performances, sometimes called “pandering,” rather than more defensible substantial performances such as fostering student learning in the classroom, enabling home support of learning, and providing important information for parents as legitimate stakeholders (Epstein, 1985; Peterson, 1989a, 2000; Scriven, 1973, 1988). However, the multiple authors of the actual survey used in this study resulted in using some items not tested in previous studies (e.g., “Child treated as an individual”). The factor analysis of themes that underlay the literal item reports revealed an unexpectedly strong sentiment on the part of parents for Humane personal treatment of pupils (see Tables 3 and 4).

This set of surveys was computer scored. Although this practice leads to increased dollar costs per form and administrative time costs over hand
scoring, it does save logistical and record keeping costs as well as clerk time in scoring. The machine scoring of all items compares favorably with the hand scoring of only the global item by more immediately providing formative feedback information to teachers.

The three-point response scale was discriminating enough to make item analysis and comparisons possible to better understand the survey. However, some analyses, particularly internal reliability (see “alpha if deleted” in Table 4), showed a relatively low variability. A five-point scale is recommended for better discrimination both quantitatively and qualitatively as it permits a distinction between degrees of agreement and disagreement.

While the internal reliability of the survey used in this study was estimated, the important questions of consistency of recording opinion over time periods (both brief-survey/re-survey and career long patterns-annual consistency) were not documented in this study. These are important empirical issues for the future.

One important topic not addressed by this study is that of teacher attention to differences in culture, language, and physical condition among teacher, classroom, students, and parents. A future study should examine some item like: This teacher did a good job of dealing with differences in culture, language, or physical condition with my child and me.

This study offers advice to teachers who are interested in fostering positive parent ratings, both for public relations and enhanced professional performance with pupils. First, it is important to be concerned with relations to students as significant people. While this can be more difficult with large classes, or multiple classes with large numbers of students during the day, these characteristics are important for parent judgment about teacher quality. Second, visible documentation of actual pupil progress and achievement is a significant concern. Third, teachers with explicit strategies for communicating with parents tend to fare more positively in parent surveys.

Appendix A: Recommended Parent Survey Form

PARENT/GUARDIAN SURVEY

Teacher’s name ________________________________

Your child’s teacher asked for a survey of parents to make the class the best it can be. Please circle the following responses that describe your experience with the teacher. No individual parents will be identified with these survey forms. Thank you for helping!

Did you ask the teacher for: Did the teacher give you:

1. An overview of class content & goals? Yes/No Yes/No
2. Description of student’s progress? Yes/No Yes/No
3. Ideas for home support of learning? Yes/No Yes/No
For each of the following, circle the number that best describes your opinion:

4. Did your child know what was expected in this class? 5 4 3 2 1
5. Was the classroom work the right difficulty for your child? 5 4 3 2 1
6. Did the teacher treat your child with respect, care, and knowledge of child’s needs? 5 4 3 2 1
7. Were you satisfied with your child’s overall school experience as provided by this teacher? 5 4 3 2 1

Do you have any comments for the teacher?: Yes Somewhat No

References

House ER (ed) (1973) School evaluation: the politics and process. McCutchan, Berkeley, CA
Lawton M (1997) Parents in N.Y. district to critique teachers. Education Week, September 17, p 3


Scriven M (1973) Handbook for model training program in qualitative educational evaluation. University of California, Berkeley, CA


Waller W (1932) The sociology of teaching. Wiley & Sons, NY