

In the Eye of the Beholder:

Teachers as Perceivers

Rena D. Harold, Jacquelynne S. Eccles, Janis Jacobs,

Allan Wigfield, Phyllis Blumenfeld, and Amy Aberbach

The University of Michigan

Paper presented at the annual meeting of the American Education Research Association, San Francisco, March, 1989.

Teachers have ample opportunity to observe and evaluate students over the course of their daily interactions with them. Researchers have long been interested in the accuracy and distinctiveness of teachers' perceptions (e.g., Brophy & Good, 1974; Brophy & Everston, 1981; Cooper, 1979; Dusek, 1985; Eccles & Wigfield, 1985; Rist, 1970). Most of these researchers, however, have limited their focus to the academic domain and, even within that domain, have taken a more global approach rather than focusing on particular subject areas. How discriminate and accurate are teachers' perceptions of their students across the various domains in which they have the opportunity to observe them? Are they better at judging children in academic subjects such as math and reading than in other areas like sports? How consistent are the ratings across time, and does this differ by domain or by construct? To what extent are teacher ratings of children's personalities related to their ratings of their abilities and efforts in various areas or domains? Finally, what is the relationship between teacher ratings of ability in various domains and actual aptitude measures given to children?

This paper will address these questions using data collected as part of a larger study of children, their parents and teachers. The data discussed here are from 540 of the children and their teachers. An equal number of boys and girls, drawn from kindergarten, first, and third grades, participated in this project in the spring of 1987, and again in the spring of 1988. These children attend public school in two middle income districts in primarily white suburbs of a large midwestern metropolitan city.

Teachers completed individual assessment questionnaires on each participating child. They were asked to rate their students using Likert-type scales on a variety of personality characteristics as well as their perceptions of the children in the following domains: math, reading, social, sports, music,

and art. For each domain, they were asked about the following constructs: talent, effort, importance of quality performance to the child, the extent to which the child worries about doing well, the extent to which the child gives up when faced with a difficult problem, and the teacher's expectation for the child's performance during the next year.

Research Questions:

1) How discriminate are teachers' perceptions of their students across the various domains in which they have the opportunity to observe them?

Teachers' ratings were highly correlated across domains for each construct except for their ratings of the children's abilities, indicating a possible "halo" effect (Brophy & Evertson, 1981). For example, to the extent that they thought that the child felt it was important to do well in one domain, they tended to rate all the domains that way: the median inter-correlations for importance was .62. Apparently, teachers see each of the measured constructs, except innate ability, as a trait-like characteristic that transcends domain.

However, when asked how much innate ability or talent the child had in each domain, the six domains formed three clusters. The two academic areas - math and reading - were correlated at .79; the other four domains formed two clusters. Ratings of innate talent for social activities and sports were correlated at .52. This may be a result of teachers judging ability in these domains on the basis of who participates in sports, which may be seen as a more social characteristic. Ratings of innate talent in the two artistic domains, music and art, were correlated at .61. These three dyads were not related to each other. Apparently, with regard to innate ability, teachers distinguish across at least three distinct domains: academic talent, social talent, and artistic talent. These patterns of relationships did not differ across grades, nor from one year to the next. We next asked 2) How do

ratings of innate ability or talent in each domain vary by gender or by grade?

Earlier research exploring gender differences in teacher ratings have not produced consistent findings (e.g., Dusek & Joseph, 1983; Hoge & Butcher, 1984; Keogh & Smith, 1970; Prawat & Jarvis, 1980; Stevenson, Parker, Wilkinson, Hegion, & Fish, 1976). We found no significant gender differences in teachers' ability ratings of their students in math, reading, or in making friends (see Figure 1). However, boys were seen as being much more talented in sports ($F=8.97$, $p<.01$; means=4.33 for girls, 4.70 for boys), while girls are seen as having more ability in art ($F=10.05$, $p<.01$; means=4.95 for girls, 4.57 for boys), and in music ($F=13.92$, $p<.001$; means=4.70 for girls, 4.34 for boys). This is interesting given that there is really no evidence of objective gender differences in art and music, and while the percent of variance accounted for by sex on a standardized performance test in sports is approximately the same as the percent of variance accounted for by student sex on the teachers' ratings, it is small (2%).

There are no significant grade differences in teachers' perceptions of talent either at time 1 when the students are in kindergarten, first and third grades or at time 2 when they are in first, second, and fourth grades. However, girls' ability in sports is rated lower at time 2 as is boys' ability in art.

3) How consistent are the teachers' ratings across time? Does their consistency differ by construct or by domain?

Identical questions were asked of the teachers at time one and time two that allow us to compare how consistent their ratings of the same students are, on the average, from one year to the next. The data discussed here are correlational and reflect the ratings that two different teachers gave the same students. However, if the trait-like characteristics that teachers see

in children reflect real consistencies in the children themselves rather than teachers' perceptual biases, then the constructs should be fairly consistent across time. In contrast, if the reality exists in the eye of the beholder, as the person-perception literature suggests, then ratings of the same children by different teachers may not be very consistent.

For the whole sample, looking across domains, within constructs, as seen in the ability example in Figure 2, it appears that the teachers' ratings in the two academic areas, math and reading, are more highly correlated with each other from year one to year two than are ratings in the other areas. Although lower in magnitude, the correlations for sports, social, and art are also significantly greater than zero. The greater consistency in ratings of math and reading may reflect more specific performance information in these areas. A similar pattern characterizes the data for the teachers' ratings of the effort, importance, and persistence constructs.

This issue of consistency is important because there is a belief that teacher expectancies can affect interaction with the child in such a way as to create a self-fulfilling prophecy (Brophy & Good, 1974). Although much literature in the academic area (Brophy & Evertson, 1981; Keogh & Smith, 1970; Stevenson et al., 1976), and to a degree in the social area (Walker, 1967), reports that teachers' estimates are accurate and consistent, these data show that there is less consistency in the non-academic domains. As you will recall, teachers reported more marked sex differences in these domains. The low consistency suggests that teacher bias may play a greater role in teacher perception in these domains. If this is true, we need to question the validity of these reports of sex differences and look very carefully at the issue of self-fulfilling prophecy effects in these domains. These concerns have begun to be addressed by researchers in the area of sports (Martinek, 1981).

4) How do teachers rate personality, and what is the relationship among ratings of personality variables and the teachers' perceptions of ability and effort across domains?

A factor analysis of the personality characteristics that the teachers rated yielded six factors. These factors were obtained at each of the three grades, and across both waves of data collection. Based on these factors, the following scales were formed: anxious, leader, pro-social, creative, responsible, and confident. Cronbach's coefficient alphas were calculated to test the reliability of the scales; they ranged from .83 to .87. Mean ratings are shown in Figure 3 summarizing teacher ratings for boys and girls at waves 1 and 2. There were significant gender differences only in their ratings of prosocial and responsible characteristics, with girls being rated as possessing more of both ($F=20.94$, $p<.001$; means=27.88 for girls, 25.55 for boys and $F=24.00$, $p<.001$; means=21.80 for girls, 19.56 for boys respectively). This is consistent with previous research that has found teachers to rate girls more positively in general than boys (Baum, Brophy, Evertson, Anderson & Crawford, 1975; Brophy & Evertson, 1981; Stevenson et al., 1976). There were no significant differences between years. There was, however, a significant difference by grade in year two in the teachers' ratings of anxiety ($F=9.87$, $p<.001$) and confidence ($F=3.37$, $p<.05$) as illustrated in Figure 4. Ratings of the children's anxiety increased while ratings of their confidence decreased from grades 1 to 4.

There were significant relationships across all the domains and constructs with the five "positive" personality scales. For example, teachers' views of the children's personalities were related to how able and effortful they judged the children to be in that all of the positive characteristics, i.e., leader, prosocial, creative, responsible, and confident, were related to the teachers' ratings of the children's talent and effort in each domain (see Table

1). The teachers did see certain characteristics as being more related to effort, while others were more related to ability. Table 1 illustrates that in the academic areas, prosocial characteristics are seen as being more related to teachers' views of effort than ability. In the social domain, however, prosocial characteristics are seen as being more related to ability than effort.

5) How accurate are teachers' perceptions of their students across the various domains in which they have the opportunity to observe them? What is the relationship of teacher ratings to actual performance, skill, or talent? Are teachers more accurate in some domains than in others?

Students were individually given a battery of cognitive measures developed by Stevenson, a shortened form of the Bruininks test of physical skills, and the Slosson I.Q. test. As Figure 6 shows, teachers' ratings of their students' abilities in math and reading were most significantly related to the children's I.Q. scores. Several of the cognitive measures were also related: following directions and writing one's name correlated with ability in both academic areas as did writing numbers -- all skills important to academic progress. In addition, the children's score on the spatial skills measure correlated with their teacher's ratings of their math ability, as did their score on the physical skills test which involves both large and small motor tasks such as copying a picture of pencils. This is an interesting relationship as small motor skill is important, for example, in drawing figures for math problems, and ratings of coordination have been positively linked to academic achievement (Brophy & Evertson, 1981; Stevenson et al., 1976).

Ability ratings in sports were significantly related to the total Bruininks score and to some of its sub-components, particularly the large motor skills such as running and broad jump. Teachers reported to us that they were much more confident of their ratings of their students' abilities in the two areas in which they had the most contact with the children, i.e., math

and reading. They were hesitant to rate the children in the other domains. However, these data show that the teacher's ratings of the children's abilities are closely related to the children's actual performance on aptitude measures tapping math, reading, and sports. These results suggest both good and bad news. Teachers appear to feel most confident and be most accurate in the areas in which they have the most experience, and in which the self-fulfilling prophecy effect may not be as biasing as we had first thought. What we do not know, however, is whether the degree of confidence teachers have in their ability ratings affects their own tendencies to act on these ratings. This is a particularly important question to answer in the non-academic domains where the extent to which teacher expectations may affect the development of children's talent in these areas is unknown.

References

- Baum, M., Brophy, J., Evertson, C., Anderson, L., and Crawford, J. (1975). Grade and sex differences in adjective descriptions of elementary school children given by their teachers and by classroom observers. Report No. 4033, Research and Development Center for Teacher Education, University of Texas, Austin (ERIC ED 146 154)
- Brophy, J., and Good, T. (1974). Teacher-student relationships: Causes and Consequences. New York: Holt, Rinehart and Winston.
- Brophy, J., and Evertson, C. (1981). Student characteristics and teaching. New York: Longman, Inc.
- Cooper, H. M. (1979). Pygmalion grows up: A model for teacher expectation communication and performance influence. Review of Educational Research, 49(3), 389-410.
- Dusek, J. B., & Joseph, G., (1983). The bases of teacher expectancies: A meta-analysis. Journal of Educational Psychology, 75(3), 327-346.
- Dusek, J. (Ed). (1985). Teacher expectancies. Hillsdale, NJ: Erlbaum.
- Eccles, J., & Wigfield, A., (1985). Teacher expectations and student motivation. In J. Dusek (Ed.), Teacher expectations. Hillsdale, NJ: Erlbaum.
- Hoge, R., & Butcher, R. (1984). Analysis of teacher judgments of pupil achievement levels. Journal of Educational Psychology, 76, 777-781.
- Keogh, B., & Smith, C. (1970). Early identification of educationally high potential and high risk children. Journal of School Psychology, 8(4), 285-290.
- Martinek, T. (1981). Pygmalion in the gym: A model for the communication of teacher expectations in physical education. Research Quarterly for Exercise and Sport, 52(1), 58-67.
- Prawatt, R., and Jarvis, R. (1980). Gender difference as a factor in teachers' perceptions of students. Journal of Educational Psychology, 72, 743-749.
- Rist, R. C. (1970). Student social class and teacher expectations: The self-fulfilling prophesy in ghetto education. Harvard Educational Review, 40, 411-451.
- Stevenson, H., Parker, P., Wilkinson, A., Hegion, A., & Fish, E. (1976). Predictive value of teachers' ratings of young children. Journal of Educational Psychology, 68, 507-517.
- Walker, R. (1967). Some temperament traits in children as viewed by their peers. Monographs of Society for Research in Child Development, 32, 1-36.

TEACHERS' RATINGS OF CHILDREN'S ABILITIES

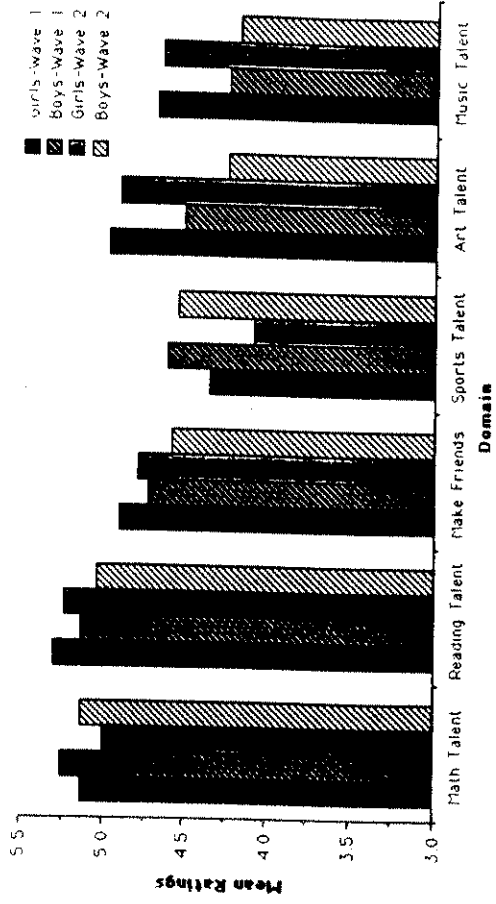


Figure 1

RELATION OF TIME 1 & 2 TEACHERS' ABILITY RATINGS

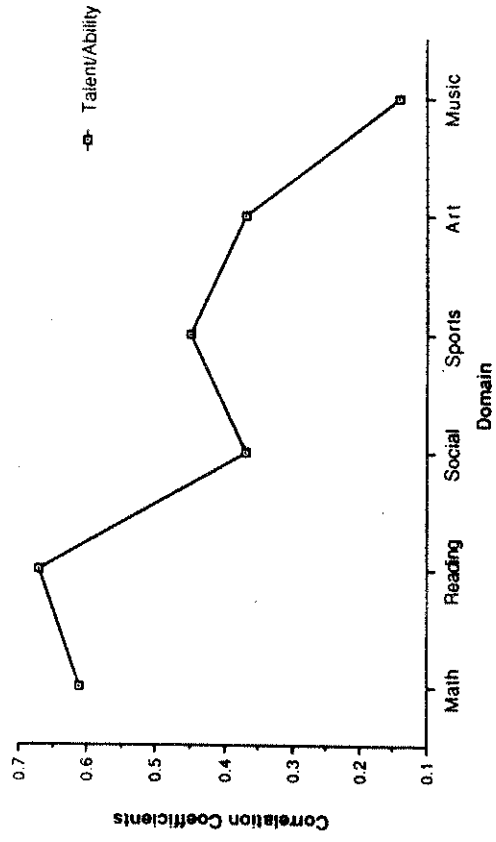


Figure 2

TEACHERS' RATINGS OF CHILDREN'S PERSONALITIES

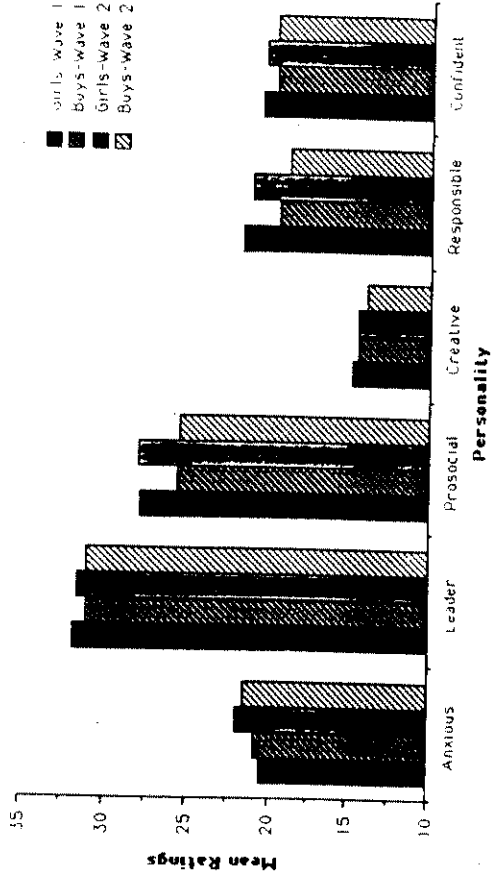


Figure 3

TEACHERS' RATINGS OF CHILDREN'S ANXIETY & CONFIDENCE BY GRADE

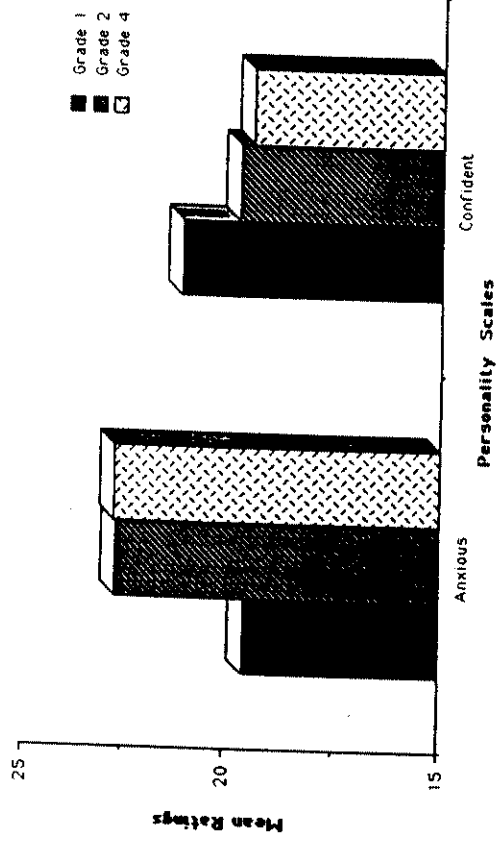


Figure 4

RELATION BETWEEN TEACHERS' RATINGS OF PROSOCIAL AND ABILITY AND EFFORT

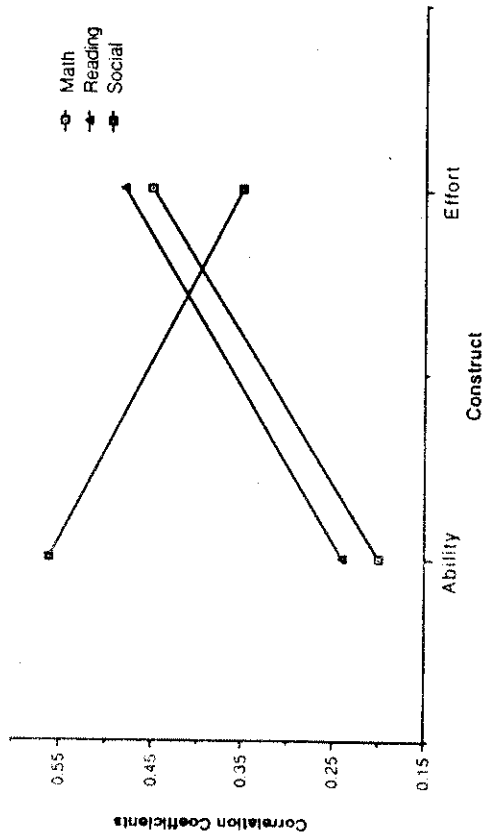


Figure 3

Table 1

Relation between teachers' ratings of the children's personality and the children's efforts and abilities: first, second, and fourth graders.

	Leader		Prosocial		Creative		Responsible		Confident	
	A	E	A	E	A	E	A	E	A	E
Math	.56	.55	.28	.52	.59	.43	.44	.66	.50	.43
Reading	.58	.54	.21	.54	.63	.45	.44	.68	.52	.44
Social	.60	.49	.61	.59	.46	.36	.41	.38	.52	.42
Sports	.46	.45	XX	.29	.37	.36	.22	.33	.39	.36
Art	.38	.36	.21	.37	.53	.47	.28	.39	.36	.35
Music	.47	.47	XX	.33	.43	.40	.32	.44	.38	.38

Notes. "A" represents the teachers' ability ratings; "E" represents the effort ratings. The zero order correlations listed in this table are all significant at p<.001.

RELATION OF TEACHERS' ABILITY RATINGS TO APTITUDE TESTS

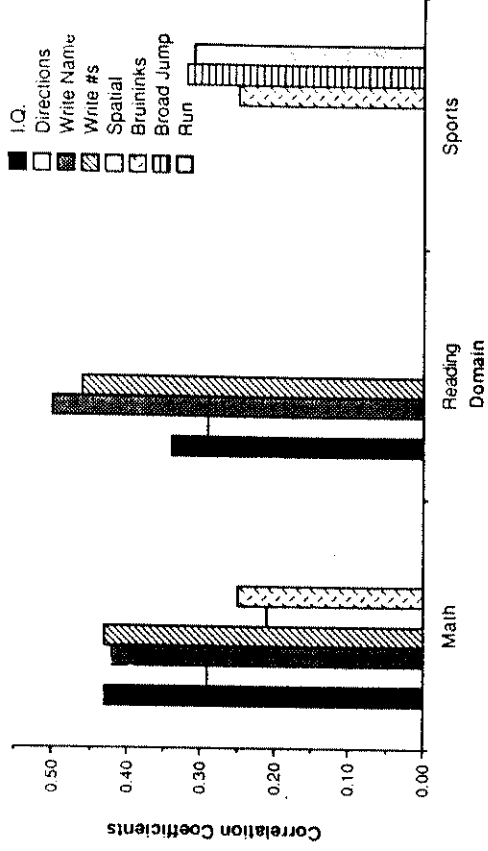


Figure 6