Behavioral Deviance and Teacher Ratings of Prosocial Behavior

Preliminary Findings


Abstract. The development of a new 'Teachers' Questionnaire concerned with rating children's prosocial behavior is described. Validation was achieved by comparing questionnaire scores with independent peer judgment of prosocial behavior. Test-retest reliability was satisfactory (r = +0.81). In a preliminary study of 132 8-year-old children, those children rated as showing behavior deviance of the conduct disorder type had significantly lower prosocial scores than children rated as neurotic or undisturbed. Implications of this finding and areas of future research are discussed.


Over the past 15 years increasing interest has been shown in the moral development of children and in the prosocial behavior of children and adults. The reasons for this are probably complex. One stimulus seems to have been the notable absence of prosocial behavior by the public at large in certain notorious incidents (Rosenthal, 1964). Second, there has been the emergence of the theory in moral developmental psychology which suggests that altruistic behavior may be an independent and biologically valuable drive, whereas previous theories, particularly psychoanalytic ones (based on studies of abnormal individuals), had stressed the development of altruism as an ego defense against primarily selfish and antisocial biological drives (Hoffman, 1975). During the same period biologists such as Edward Wilson (1976) have taken an increased interest in altruistic behavior occurring in animals.
Prosocial behavior is an umbrella term which covers a number of similar, but not necessarily related, social actions such as helpfulness, generosity, and cooperation. The theoretical issues raised have been well discussed by Hoffman (1975), while more recently Mussen and Eisenberg-Berg (1977) have provided a coherent review of the information which has accumulated as a result of experimental and quasi-experimental work with children. Briefly, the evidence suggests that a child's propensity to perform a prosocial act increases with age up to middle childhood and that this is associated with developmental changes in moral reasoning and role-taking ability. Individual differences resulting from socialization experiences are thought to influence prosocial behavior, and there is some evidence of a persistent prosocial character trait. Sex differences have not been consistently found. A number of situational variables are thought to influence the tendency to perform prosocially, such as mood, the behavior of others in the vicinity, and the attractiveness of the person to be helped.

Systematic observation of individuals' prosocial behavior in truly natural surroundings has been virtually absent, mainly because the technique is time consuming and difficult. This is, however, an important gap because no matter how subtle the strategy used in experimentally contrived situations in a natural setting, one may still be measuring children's knowledge of how they ought to behave rather than how they really behave. A few investigators (Friedrich and Stein, 1973; Murphy, 1937) have attempted measurement of naturally occurring behavior in children, using observers' ratings of such items as "cooperativeness," "generosity to others," and "sympathy." Such observations, however, have been limited to nursery-age children.

To improve the assessment of naturally occurring prosocial behavior there is a need for an easily used, reliable instrument measuring spontaneous prosocial behavior in older children.

We have chosen the school as our natural setting and here report the preliminary development and validation of a questionnaire of children's prosocial behavior for use by teachers. The link between prosocial behavior and psychiatric disorder has never been systematically examined in population studies. Despite limitations in interpretation due to uncertainty regarding the motivation of such qualities as generosity and helpfulness, it seemed to us worthwhile to investigate the association between the two. We have also examined, therefore, the association between prosocial behavior and behavioral deviance.
Method

Following introductory discussion, the heads and deputy heads of five junior schools in an inner London borough were asked to compile lists of actually observed behavior in which 8-year-old children in the course of the school day were being helpful, generous, considerate, or cooperative. In the preparation of a final list of items for the questionnaire from the lists provided by the teachers, it was possible to condense some behaviors which were similar in character into one item and to eliminate those behaviors which were uncommon or peculiar to a particular school or class. For example, one item, “offers to help other children who are having difficulty with a task in the classroom,” was derived from a number of observations of children spontaneously helping others who had spelling, reading, drawing, or other practical difficulties. A striking example of altruism occurred when a whole class of 7- to 8-year-olds spontaneously, and without the teacher’s knowledge, made a collection of money as a leaving present to a boy from a children’s home who was moving to another home far away from the school, but here the circumstances were too uncommon to form the basis of an item.

Teachers were also aware that the motivation behind prosocial acts varied considerably, from the genuinely unselfish to the obsequiously compliant. As far as possible items were selected in which the motivation appeared to be genuinely prosocial (i.e., unselfish and not serving the neurotic needs of the individual). However, it should be stressed that the questionnaire is designed only as a measure of observable behaviors and not as an instrument for investigating the intention behind the behavior, the antecedents of any prosocial acts, or the desired consequences.

The questions and teachers’ instructions for the questionnaire are shown in Table 1. Teachers rated each behavior on a 3-point scale: “certainly applies,” “applies somewhat,” and “doesn’t apply”; the categories were given scores of 2, 1, and 0 respectively. Five classes of 8-year-olds were selected. For each pupil background data on age and sex were collected, and teachers were asked to score the prosocial questionnaire for each child. The results were compared with those obtained by the simultaneous administration of 16 prosocial items from the Pittsburgh Adjustment Survey Scale (PASS) described by Ross et al. (1965), a questionnaire whose reliability and validity have been tested by Miller (1972). Teachers were
then asked to complete a questionnaire designed to identify those children most likely to be suffering from a behavioral disturbance (Rutter, 1967). Children classified as behaviorally deviant on this questionnaire can be subdivided into those with antisocial deviance, neurotic deviance, and mixed deviance on the basis of scores on individual items (Rutter et al., 1970).

In an effort to obtain concurrent validation of the prosocial behavior questionnaire (PBQ), we sought pupils' opinions about each other's prosocial behavior, using a modification of the class play technique described by Bower (1960). The class play technique was originally designed as an instrument for obtaining peer ratings of emotional disturbance; Cowen et al. (1973) found it to be highly predictive of later referral to a psychiatric clinic. A number of other studies have suggested that pupils' opinions form a good basis for predicting later psychiatric disturbance or delinquency (West and Farrington, 1973). It might be expected that pupils' opinions will be just as valid when used to judge prosocial behavior in the class or playground. In this modification of the Class Play Technique, each child in a class is provided with a sheet of paper on which are instructions followed by a list of 14 observable items of prosocial behavior (which are 14 slightly rephrased versions of questions from the teachers' prosocial questionnaire) and are asked to choose from their class one boy and one girl who would be best suited to play that particular role in a play put on by the class. It is then possible to count the number of times each child is chosen for each role by the class as a whole, and a total score for each child is obtained by adding together the number of times he or she was chosen for each of the 14 roles. The total score for each child is then corrected to allow for variations in the total number of children in each class, and the different boy/girl ratios in each class. This correction allows comparisons to be made using the Class Play totals for all the classes pooled together. The rank-order correlation between the score obtained on the teachers' prosocial questionnaire and the corrected total score obtained from the Class Play was used as a measure of concurrent validity. An additional measure of concurrent validity was obtained from the correlation between the child's score on the PBQ and the score on the PASS prosocial items. The Class Play was administered to three classes. Because of limited time it was not possible to perform the methodologically superior validation by an independent trained observer.
In this age group the pupils spend a majority of their time with one teacher, and it is not possible to obtain interrater reliability easily. However, test-retest reliability with the same teacher at an interval of two months was obtained in two classes.

**RESULTS**

*Administering the PBQ*

It was soon apparent that the questions demanded of the teacher a thorough knowledge of each child, so that the instruction to the teachers that they should only rate children they have known for more than two months (a school term) probably greatly increased the reliability of the questionnaire.

The experimental version of the PBQ is reproduced in table 1. For the 73 boys the mean total score on the PBQ was 16.18 (SD = 9.74), and for the 59 girls 19.08 (SD = 9.32); this difference is not statistically significant. The distributions of the total PBQ scores for boys and girls were also similar.

Item to partial-scale correlations for boys varied within the range +0.33 to +0.78, for girls from +0.23 to +0.76, for girls and boys combined from +0.38 to +0.75. In all cases the modal range was +0.60 to +0.69, and in all but two items the correlation was significant at the 0.001 level (items 17 and 18, p < 0.05).

A principal component analysis of the correlation matrix between each item on the questionnaire revealed a first component that explained 41% of the variance. There were three further components with eigenvalues greater than one explaining 10.1%, 8%, and 6.6% of the variance respectively.

*Validity and Reliability*

The results of the two tests of validity are shown in table 2. Sexist discrimination in the Class Play was controlled for by asking each child to nominate a boy and a girl for each role, thus eliminating the social dominance of boys. The peer ratings of prosocial behaviors obtained from the Class Play method correlated highly with the ratings obtained from the PBQ.

The second test of validity was to compare scores obtained on the PBQ for each child with scores obtained for the same children on the prosocial subscale of the PASS. The results shown in table 2 were satisfactory. Test-retest correlations for the PBQ at two months are also shown in table 2.
Table 1

TEACHERS' QUESTIONNAIRE
(Experimental Version)

Below is a list of behaviors which may be shown by a child during the school day. On your knowledge of the child over the last two months could you place a cross in the appropriate box.\(^1\) If the child definitely shows the behavior described by the statement, place the cross in the box under “certainly applies.” If the child shows the behavior but to a lesser degree or less often, place the cross under “applies somewhat.” If as far as you know the child does not show the behavior, place the cross in the box under “doesn’t apply.” Place ONE cross against EACH statement.

<table>
<thead>
<tr>
<th>Doesn’t Apply</th>
<th>Applies Somewhat</th>
<th>Certainly Applies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. If there is a fight or a quarrel, will try to stop it.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Will invite bystanders to join in a game.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Goes to the help of someone who has been hurt.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Helps to keep other children quiet in class or assembly.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Is considerate of the teacher’s feeling.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Shares out sweets or extra food at lunchtime.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Tries to be fair in games.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Takes the opportunity to praise the work of less able children.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. When choosing partners for an activity often chooses someone who might otherwise be left out.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Is generous in contribution toward gifts and charities.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Will offer to show a new child around the school.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Offers to help other children who are having difficulty with a task in the classroom.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Shows concern for the welfare of younger children in the school when teachers are not present, for example, in the playground at lunchtime.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Offers to help children who are feeling sick.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Can work easily in a small peer group.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Doesn’t need reminding if asked to carry out a regular task, such as helping with the school milk, or feeding the class pet.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Settles down to work quickly.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Looks embarrassed if someone else in the classroom makes a mistake.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Will clap or smile if someone else does something well in class.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. Volunteers to help clear up a mess someone else has made.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^1\) The original had separate boxes drawn under each statement.
Table 2

The Validity and Reliability of the PBQ (Correlations of the PBQ Total Score with Other Measures: Boys and Girls Combined)

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Correlation</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Play Score (peer rating)</td>
<td>77</td>
<td>+0.50</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>PASS prosocial score</td>
<td>60</td>
<td>+0.77</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>Test-retest of PBQ at 2 mths.</td>
<td>60</td>
<td>+0.81</td>
<td>p &lt; .001</td>
</tr>
</tbody>
</table>

Age and Sex

The children in the sample were all in the age range between 7 years, 6 months and 8 years, 11 months; it was therefore not possible to look at any effect which age might have upon scores on the PBQ. We found no significant sex differences on any of the measures of prosocial behavior.

Behavior Deviance

There was a correlation of −0.45 (p < 0.001) between the total score on the Rutter Teachers’ Questionnaire of behavioral deviance and the total score on the PBQ. Children identified as being at higher risk of psychiatric disorder on the Rutter Questionnaire (Score 9+) were allocated either to subgroup of Neurotic Disorder, or Conduct plus Mixed Disorder, using the method described by Rutter et al. (1970). The total PBQ scores for these subgroups and for the nondeviant group are shown in table 3. They reveal an unexpected difference. The Neurotic subgroup

Table 3

Mean Total Scores on the PBQ for Behaviorally Deviant and Nondeviant Groups

<table>
<thead>
<tr>
<th></th>
<th>Boys</th>
<th>Girls</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>1. Nondeviant group</td>
<td>55</td>
<td>18.00</td>
<td>9.26</td>
</tr>
<tr>
<td>2. Neurotic deviant</td>
<td>7</td>
<td>13.29</td>
<td>8.75</td>
</tr>
<tr>
<td>3. Conduct and mixed deviant groups</td>
<td>11</td>
<td>6.46</td>
<td>5.96</td>
</tr>
</tbody>
</table>

ANOVA

F = 7.88
p < .001

P = 9.71
p < .001

p = 17.27
p < .001

Sheffe’s Multiple Range Test

Group 3 2 1
Group 3 2 1
Group 3 2 1

* Groups with means not significantly different from one another at p .05 are given a common underlining.
scored only very slightly lower on the average than the nondeviant group, whereas the Conduct plus Mixed subgroup had markedly lower PBQ total scores. An analysis of variance and subsequent Scheffé's Multiple Range Test showed that for boys and girls separately the mean total score of the Conduct plus Mixed subgroup was significantly lower (p < 0.001) than the mean total score of the nondeviant group, but not significantly lower than the mean total scores for the Neurotic group. However, when the scores for boys and girls are combined, there is again no significant difference between the mean total scores between nondeviant and Neurotic groups, but the Conduct plus Mixed subgroup mean total score is significantly lower (p < 0.001) than both the nondeviant group and the Neurotic subgroup (table 3).

**DISCUSSION**

The results suggest that the PBQ has concurrent validity as a measure of the prosocial behavior shown by 8-year-old children in school. Its results correlate highly with those obtained from the prosocial items on the PASS Teachers' Questionnaire. The peer ratings on children's prosocial behavior using the Class Play method also correlated significantly with the total PBQ scores.

The Class Play method of obtaining peer opinions of prosocial behavior could, however, be influenced by a number of other factors such as popularity, intelligence, or attractiveness (Gronlund, 1959), which might also affect teacher ratings. Thus validation by direct observation is also necessary, and is currently being attempted.

The correlations between the scores on individual items and the total prosocial scores provided some evidence of construct validity, though Rushton (1976) estimated that the average correlation between items of prosocial behavior shown in experimental settings was as low as +0.3.

The distribution of total prosocial scores for boys and girls followed an approximately normal distribution curve, which was thought to be good evidence of the ability of the PBQ to quantify prosocial behavior along a continuum. The lack of difference in prosocial behavior shown by boys and girls is not unexpected in view of the conflicting results reported in previous studies. It is, however, possible that sex differences do occur in different age groups.

It is of considerable interest that children with an antisocial form of deviance had such a low mean score of prosocial behavior. It is
first necessary to emphasize that the presence of antisocial behavior by no means excludes the presence of helpfulness and other positive traits. Prosocial behavior is not the necessary obverse of antisocial behavior. Indeed, it has sometimes been suggested that aggressive children who truant and steal are likely to adhere to the values of their peers especially closely, and one might therefore have predicted at least average prosocial scores for the group. The finding needs further investigation. First, a possible negative halo effect by the teachers making the ratings needs exclusion, and a validation of the PBQ by direct observation which is currently being undertaken should to some degree clarify the issue. Teacher ratings might be influenced by a variety of other characteristics of the children such as academic ability, attainment, family background, and social class, together with other personality traits such as extroversion. Various factors, however, make it likely that a meaningful association exists. The written instructions for the PBQ do draw attention to the need to rate only actually observed behavior, and the questions are worded as far as possible to gain information about the exhibited behavior rather than teacher judgments about character or personality. Children with antisocial deviance were rated in prosocial behavior not only by teachers, but also by their peers using the Class Play Technique. Other work (Chandler, 1973) suggests that training to improve the capacity of delinquent boys to see the world from the perspective of others (a skill necessary for effective helpfulness) actually decreases the rate of subsequent delinquency when compared with placebo treatment of a similar type. By contrast, neurotic deviance might have been expected to be associated with either a high or a low mean score. Inhibited children might have been expected to show lower scores because prosocial behavior requires a certain degree of assertiveness for its execution. Alternatively, the overconscientiousness sometimes associated with neurotic symptoms might have been expected to be linked with compulsive dutifulness and a high mean score. Neither of these was found, nor was there a wider than expected dispersion of scores in the neurotically deviant group—a finding that might have been expected if both these factors had been operating but counterbalancing each other. Again the finding needs replication on a larger group of children. A further possibility is that in rating the neurotic children the teachers may have been operating but counterbalancing each other. Again the finding needs replication on a larger group of children. A further pos-
The development of a reliable and valid method for assessing prosocial behavior in populations of children would allow systematic investigation of the factors operating in different populations which augment or diminish its occurrence. With this in mind we intend to develop the method described here.

Finally, we should like to point to an unexpected side effect of our inquiries. After completing the questionnaire, teachers often commented that answering the questions about prosocial behavior made them think about their pupils in a new way. They were not used to bringing to mind children’s good behavior, and found the exercise stimulating and enlightening. The fact that a group of psychiatrists who might have been expected to focus solely on troublesome or troubled behavior were interested in learning about prosocial behavior must surely have had some influence here. We would like to think that a small switch of professional psychiatric and psychological attention to these positive aspects of behavior might achieve at least a slight increase in the likelihood of their occurrence.

REFERENCES


