

Measuring Children's Aggression With Teachers' Predictions of Peer Nominations

L. Rowell Huesmann, Leonard D. Eron, Nancy G. Guerra, and Victoria B. Crawshaw

Peer nominations have been used very successfully to assess aggressive dispositions in children but are costly to administer in situations in which a subject population is scattered among many classrooms. In the present study, the authors evaluated an alternative measure, the Teacher Prediction of Peer-Nominated Aggression. This measure proved to be highly reliable and valid—a better predictor of peer nominations of aggression than teacher checklist ratings of aggression. The teachers' predictions of peer-nominated aggression also displayed the same pattern of interrelations and gender differences as actual peer nominations of aggression. Finally, the teachers' predictions of aggression were more accurate than their predictions of other behaviors.

One problem that has beset research on personality is the difficulty of directly observing and accurately measuring personality traits as they are manifest outside the laboratory in real life. Traits cannot be directly observed; they can only be inferred from observed behavioral dispositions. In fact, traits are best viewed simply as behavioral dispositions.

To measure a behavioral disposition, the researcher must develop an operational definition of the construct of interest in terms of observable behaviors. The researcher must observe these behaviors directly or obtain judgments from individuals who have had the opportunity to observe them in the past. These observations must then be counted, rated, subjected to arithmetic operations, and perhaps processed by computer programs designed to detect underlying latent structure. The results of such analyses must then be interpreted to yield a plausible measurement scale for the underlying disposition. There are many potential pitfalls on this path from behavioral observations to trait inference. Generally, researchers are on firmer ground if a particular trait of interest can be connected very closely with observable behaviors.

One of the most widely researched personality traits, at least since the late 1930s, has been aggression. Although individuals vary widely in their characteristic levels of aggressiveness, ex-

treme aggression at times falls in the realm of psychopathology and has been so considered by some researchers (Kazdin, 1987). Aside from the fact that extreme aggression is of concern in its own right, it often co-occurs with other psychopathologies. A number of longitudinal studies have shown that early extreme aggressiveness, in fact, is predictive of a range of later psychopathologies (Farrington, 1991; Huesmann, Eron, Lefkowitz, & Walder, 1984; Offord, Boyle, & Racine, 1991).

Aggressiveness has also been a popular disposition for study because it can be closely linked to observed behavior. An aggressive behavior has generally been defined as a behavior that is intended to injure or irritate another person (Eron, Walder, & Lefkowitz, 1971). Aggressiveness, then, is the disposition to engage frequently in behaviors that are intended to injure or irritate another person. The one difficulty this definition presents for measurement is the intentionality component. Whether or not an observed behavior injures or irritates another person can usually be determined without much difficulty, but the intention behind the behavior may be more difficult to divine, particularly when one is dealing with children. Self-reports do not solve the problem, either, as even the individual behaver may not be aware of the intentions behind the behavior (Nisbett & Wilson, 1977). The solution to this problem that has been widely adopted is to infer intentionality when behaviors that injure or irritate another become frequent. Although such a solution may not help to classify any one behavior, it is sufficient for a classification of individuals' dispositions to behave aggressively, as long as a sufficient number of observations can be made to assess frequency accurately. The majority of measurement techniques that have been derived to assess individuals' aggressive dispositions outside of the laboratory have adopted this approach.

For assessing aggression in children, one of the most popular techniques has been peer nominations. This procedure has the advantage of providing multiple observations of the same behaviors by a number of different observers (e.g., all the children in a given child's classroom). One such widely used procedure, the Peer Nomination Inventory developed by Eron and his colleagues (Eron et al., 1971; Walder, Abelson, Eron, Banta, & Laulicht, 1961), has been shown to be a highly reliable and valid

L. Rowell Huesmann and Leonard D. Eron, Research Center for Group Dynamics, Institute for Social Research, University of Michigan; Nancy G. Guerra and Victoria B. Crawshaw, Department of Psychology, University of Illinois at Chicago.

This research has been supported by grants from the National Institute of Mental Health and the Centers for Disease Control and Prevention. A modified version of this article was presented in July 1992 at an invited symposium on human aggression at the 25th International Congress of Psychology, Brussels, Belgium. We would like to thank Ed Czilli, Crystal Scott, and Arnaldo Zelli for their contributions to the data analyses and Sharon Grant for her contribution to the data collection.

Correspondence concerning this article should be addressed to L. Rowell Huesmann, Research Center for Group Dynamics, Institute for Social Research, University of Michigan, Ann Arbor, Michigan 48106-1248.

tool for assessing aggressive behavior in school-age children. This measure was developed over 35 years ago and is still in use today with no diminution in its relevance for children of varying ethnicity, race, socioeconomic status, or geographical location around the world (Guerra, Huesmann, Tolan, Van Acker, & Eron, 1994; Huesmann & Eron, 1986; Pulkkinen, 1987). Its concurrent validity has been demonstrated in both laboratory (Williams, Meyerson, Eron, & Semler, 1967) and field studies (Eron et al., 1971); its construct validity has been demonstrated in the many theoretically predicted relations into which it enters (e.g., Eron, Laulicht, Walder, Farber, & Spiegel, 1961); and its predictive validity has been demonstrated over 3 years (Huesmann & Eron, 1986), 11 years (Lefkowitz, Eron, Walder, & Huesmann, 1977), and 22 years, as seen in its relation to the occurrence of aggression, violence, delinquency, and criminal behavior over those periods (Huesmann et al., 1984). A number of similar peer-nomination scales have been derived and are also in widespread use in aggression research (Bjorkqvist, Lagerspetz, & Kaukiainen, 1992; Dodge & Coie, 1987; Perry, Perry, & Rasmussen, 1986; Wiggins & Winder, 1961).

A major problem with peer nominations, however, is that one must be able to identify and interview a substantial body of peers who know the target person well. As argued earlier, the validity of the frequency of a behavioral approach to measuring a personality trait like aggression depends on the extent to which a substantial sample of unbiased observations can be obtained. One of the reasons why peer nominations have proven highly reliable and valid is undoubtedly because they are based on observations over a long period of time (the school year to date) by a substantial number of raters (each classmate). Furthermore, no single rater can influence the target child's final score more than any other because each rater only nominates or does not nominate. Raters do not assign rating numbers on a scale. Obtaining a substantial sample of raters is easy in school settings when one is mostly interested in obtaining scores for the same children who are serving as raters. However, it becomes very difficult outside of school settings or when the children of interest are spread over a large number of schools or classes. It is especially problematic in longitudinal studies, as settings and acquaintance networks can change swiftly and dramatically. Thus, it is necessary to have a number of alternative methodologies for measuring the behavior of interest, in this case, aggression.

The most obvious alternative would be self-ratings. However, this has proven not to be a good substitute for peer nominations, especially when used with young children (Leibowitz, 1968). The relation of peer nominations to self-ratings (used with early versions of the peer-nomination procedures) was only .33, and thus its use as a valid measure of aggression seemed problematic (Eron et al., 1971). In more recent studies with inner-city African American children, the correlations between self-ratings and peer nominations have been even lower, that is, under .20 (Huesmann et al., 1989). It may be that the self-incriminating nature of self-ratings is too difficult to overcome when used with young children, even with assurances of anonymity. It is probably also true that young children have difficulty taking the perspective of an outside observer when judging their own behavior. Additionally, as some have argued (e.g., Caprara, 1992), children's self-ratings may be excessively biased by the views ex-

pressed to them about them by a few significant others in their life.

Another reasonable option might be parent ratings. However, the correlations between peer nominations and parent ratings have been even lower than for self-ratings, for example, .22 for fathers and .16 for mothers (Eron et al., 1971). Parents, it seems, are not good judges of how their children interact with other children outside of the home. Aggression, like other behavioral dispositions, clearly has some situational specificity. The characteristic mode of behavior of a child outside of the home may be quite different than the mode inside the home.

Another alternative is teacher ratings. When comparing the utility of various methods, it is essential that the different methods agree on what aspects of child aggression and conduct problems are measured. A close examination of the different teacher and peer measurement instruments reveals different conceptualizations (and operationalizations) of the aggression construct. Teachers and peers may be reacting to and judging different classes of behaviors and events, despite the fact that all behaviors and events may ultimately relate to conduct problems in classroom and school situations. Both peers and teachers are, in principle, social observers of the same events and behaviors, and therefore should agree. However, peers very likely form their judgments about any one of their classmates not only on the basis of personal experience, but also on the basis of the ways in which others interact with and respond to the child. Similarly, teachers may also partially form their judgments about any one of the children on the basis of how the class of children, as a whole, respond to that child. Therefore, it is reasonable to expect that teachers may be capable of accurately perceiving how the children in their classes interact with each other.

The results with previous studies of teacher ratings have been mixed, however (Achenbach & Edelbrock, 1981, 1983; Ollendick, Oswald, & Francis, 1989). For example, in a large sample of third graders studied in 1960, we obtained a rating of aggression for each child from each child's teacher, as well as from the child's peers. Depending on the set of questions used, the correlations ranged from .41 to .63 (Eron et al., 1971). Although the correlations were significant and higher than the correlations obtained between peer and self-reports or peer and parent ratings, they were not consistently high enough to justify substituting teacher nominations for peer nominations. Another problem with teachers' checklist ratings is that they generally tend to be highly skewed—even more skewed than the actual distribution of aggressive behavior across individuals (Guerra et al., 1994).

A major problem with any of the alternative techniques already discussed is that the score for each child is based on only one observer's judgment—the child, a parent, or a teacher. Yet, as was argued at the beginning of this article, the validity of the entire approach to measuring aggressive dispositions by way of frequency of behaviors that injure or irritate another person depends on obtaining multiple unbiased observations. Unfortunately, any solution that really uses multiple observers is likely to be as costly as peer nominations are. However, an intermediate approach that might achieve reasonable validity at lower cost would be one that asks a single informed observer to estimate what multiple observers would conclude. Individual bias is minimized by asking the observer not to make his or her own

judgment but simply to report what the judgment of a well-known group would be. The obvious application of this approach to a school situation would be to ask the teacher to estimate the judgments that children's peers would make about them. The current study investigates the validity of this approach.

Method

Participants

The participants for this study were second-, third-, fourth-, and fifth-grade elementary school children attending an inner-city school in a large midwestern city. Initially, the 9 teachers of these grades were asked to participate in the study, and 6 teachers agreed. Participants were thus drawn from these six classrooms, and each child was rated by only 1 teacher. Parental permission was obtained for 86% of the children in these classes, resulting in a final sample of 179 participants (82 boys and 97 girls). In particular, the sample was made up of 29 second graders, 32 third graders, 82 fourth graders, and 36 fifth graders. All of the children in the sample were African American and lived in a low-income housing project, with 96% reporting annual family incomes below the U.S. poverty level.

Measures

The children in this study were assessed on several measures that were intended to estimate aggression or constructs theoretically related to aggression. Both the children's peers and their teachers were used as data sources.

Peer-nominated indexes of children's aggression, prosocial behavior, popularity, rejection, hyperactivity, and victimization were obtained by administering the Peer-Nomination Inventory to all children participating in the study. The complete set of 22 questions, as modified for the teacher form, are shown in Table 1. Each child in a classroom is asked to nominate every child who fits a particular question by marking the child's name on a list of names. These scales are scored to provide percentage ratings of the number of times the child has been nominated for any given behavior by his or her peers out of the total number of times the child could have been nominated.

As mentioned previously, the Aggression scale has been widely used for over 30 years, and the procedure, its reliability, and its validity have been reported extensively (Eron et al., 1971; Huesmann & Eron, 1986). It consists of 10 questions about whether a child displays a specific aggressive behavior (e.g., "Who pushes and shoves other children"). Similarly, the Popularity scale has been used widely in these same studies. The Prosocial, Rejection, Hyperactivity, and Victimization scales are relatively new, but they have also been found to be reliable and valid (Huesmann, Guerra, & Eron, 1992).

Teachers provided two kinds of assessments of the children's aggression at about the same time the peer nominations were obtained. First, the teachers rated each child on the complete Child Behavior Checklist (Achenbach & Edelbrock, 1983), although only scores on the 39-item Aggression scale are evaluated in the current study. In addition, teachers completed the Teacher Prediction of Peer-Nomination Scales, which were developed specifically for the present study. The scores on these scales are intended to estimate, for any given child and for any given peer-nomination scale, the percentage of the child's peers who would nominate the child on the questions making up that scale. The teacher's response to each question was coded as the midpoint of the category that the teacher selected.

Two different methods of obtaining scores were developed. In one method, the *by child* method, the teachers would first predict the percentage of peers who would nominate one child, Child A, on each of the

22 peer-nominated behaviors, only 10 of which were aggression. Then the teachers would estimate the percentage of peers who would nominate another child, Child B, on each of the 22 behaviors. Then they would give predictions for Child C, and so on. This continued until all children had been rated. With the other method, the *by question* method, the teachers would predict the number of children who would nominate Child A, Child B, and so on, on Question 1. Then they would predict the number of children who would nominate Child A, Child B, and so on, on Question 2. This continued until scores had been obtained on all questions. The *by question* method is the one that most closely resembles the usual peer-nomination assessment procedure with children. However, the *by child* method has more potential utility as an assessment tool, as only the target child needs to be rated. In the present study, one-half of the teacher sample was given the *by child* method and one-half was given the *by question* method so the methods could be compared. The same 22 questions appearing in the children's Peer-Nomination Inventory were used and are shown in Table 1.

Results

The internal consistency reliability of the scales under each method is also shown in Table 1. The coefficient alpha for the new 10-item scale for Teacher Prediction of Peer-Nominated Aggression was .95 for the *by child* version and .97 for the *by question* version. These numbers are comparable with a computed internal consistency of .97 for the 10-item children's Peer-Nomination Scale in this sample. The internal consistencies of the other teacher-nomination scales are not as high as for aggression but are mostly good considering the size of the scale and are comparable to the alphas for the peer nominations.

Figure 1 displays the intercorrelations between the three scales for assessing aggression along with their distributions. As Figure 1 reveals, the distribution of aggression scores on the Teacher Prediction of Peer-Nominated Aggression was very positively skewed (.881), as were the scores on the teacher-rated CBC (1.07). The scores on actual peer nominations of aggression were less skewed (.589). Positive skewness is typical of the distribution of aggressive behavior among children, adolescents, and adults (Huesmann et al., 1984); however, the differential skewness of the teacher-based and peer-based scores suggests that peers discriminate more at the lower end of the range of aggression than do teachers.

The overall correlations among the three aggression measures shown in Figure 1 must be evaluated with respect to their individual and joint distributions also shown in Figure 1. The ellipses in the figures represent 50% confidence intervals for the data, and the lines represent quadratic regression fits. One can see that the teacher predictions of peer-nominated aggression correlate very highly with the actual peer nominations of aggression, significantly more highly than the Child Behavior Checklist Aggression scale correlates with either, $z = 3.15$, $p < .005$. The relation between the teacher predictions of peer-nominated aggression and the actual peer nominations of aggression is linear with no significant quadratic component. This correlation is particularly notable considering the theoretical limits placed on the correlation by the differential distribution of the variables and by the fact that different sources provided the peer and teacher data.

In Table 2, the correlations are broken down by gender. One can see that all the correlations are slightly (but not signifi-

Table 1
Items From Teacher Predictions of Peer Nominations

Scale item	Coefficient α		
	Peer method	Teacher by question	Teacher by child
What percentage of students would say that Johnny. . .			
Aggression scale	.97	.97	.95
does not obey the teacher.			
often says "Gimme that!"			
is a child who pushes or shoves other kids.			
gives dirty looks or sticks out his or her tongue at other children.			
makes up stories and lies to get other children in trouble.			
does things that bother other children.			
starts a fight over nothing.			
is always getting into trouble.			
says mean things.			
takes other children's things without asking.			
Prosocial scale	.93	.63	.89
likes to share with others.			
gets along well with others.			
helps other kids.			
does nice things to help other people.			
Popularity scale	.75	.72	.91
is someone I would like to sit next to in class.			
is a child I would like to have for my best friend.			
Rejection scale	.93	.70	.77
is a child I really don't like.			
is someone I wish were not in my class at all.			
Victimization scale	.71	.69	.48
gets picked on by other kids.			
gets hit and pushed by other kids.			
Hyperactivity scale	.82	.80	.78
gets out of his or her seat a lot.			
wiggles or moves around in his or her seat a lot.			

Note. Teachers were asked to rate what percentage of students (i.e., 0%, 1-5%, 6-10%, 11-25%, 26-50%, or over 50%) would choose each particular item to describe a particular student.

cantly) higher for boys than for girls, but the ordering of the correlations is about the same. For each gender, teacher predictions of peer nominations correlate very highly with actual peer nominations, much more highly than the CBC Aggression score correlates with peer nominations.

We also computed the correlations by grade, and no significant differences between the correlations were revealed. In Table 3, the partial correlations among the three aggression scales are presented with grade partialled out. Aggressive behavior is known to increase with grade in this age range, so it is not surprising that the correlations decrease 15% to 20%. However, the correlation between teachers' predictions and actual peer nominations is still a highly significant .66.

As the scatter plots in Figure 1 suggest, these correlations between teacher-nominated aggression and peer-nominated aggression reflect a linear relation in which predictive validity is maintained across the range of scores. For example, of those identified by their peers as being in the upper quartile for aggression, we found that 73% were also identified as being in the upper quartile by their teachers. Of those identified as being in the lower quartile by their peers, 66% were also identified as being in the lower quartile by teachers.

A regression prediction of actual peer nominations from

teachers' predictions and the children's grade and gender confirmed that grade and gender information do not add significantly to the prediction beyond the teachers' ratings. The slope and intercept for the raw regression line predicting actual peer nominations from teachers' ratings alone were .90 and .15, respectively, $F(1, 172) = 244.5, p < .0001$. These parameters reflect the significantly higher mean score for actual ratings in comparison with teachers' predictions (.27 vs. .14).

In Table 4, the correlations for the two forms of the Teacher Prediction of Peer-Nominated Aggression are compared. The correlations between actual peer nominations and teacher predictions are slightly, but not significantly, higher for the by question form than for the by child form, as was expected (.83 in comparison with .73, $z = 1.55, ns$). The by question form more nearly mimics the procedure used with the actual peer nominations. However, even this slight nonsignificant difference may be spurious. The correlations between the CBC Aggression scale and both the teacher predictions and actual peer nominations are also higher for the by question condition. There is no theoretical reason why this should be the case. The CBC procedure in fact more closely parallels the by child form. Thus, the differences between the two forms in these samples may well be attributable to sampling error.

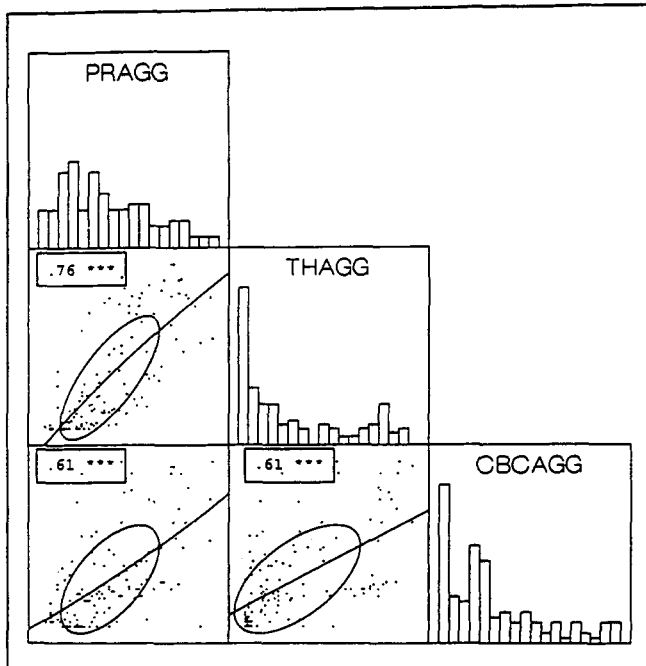


Figure 1. The frequency distributions and intercorrelations of the three aggression measures ($N = 130$). PRAGG is actual peer-nominated aggression; THAGG is teacher predictions of peer-nominated aggression; and CBCAGG is teacher-rated Child Behavior Checklist aggression. The best-fitting quadratic curve and a 50% confidence ellipse are shown for each bivariate relation. *** $p < .001$.

In Figure 2, the distributions of the other five peer-nomination scales and their correlations with teacher predictions are shown. One can see that all of the teacher scales except Prediction of Popularity are highly skewed in comparison with actual peer nominations. One can also see from the diagonal that all five of the other teacher prediction scales (Prosocial Behavior, Popularity, Rejection, Victimization, and Hyperactivity) corre-

Table 2
Correlations by Gender

Measure	1	2	3
Boys			
1. PRAGG	—		
2. THAGG ^a	.79*	—	
3. CBCAGG ^b	.62*	.63*	—
Girls			
1. PRAGG	—		
2. THAGG ^c	.69*	—	
3. CBCAGG ^b	.51*	.52*	—

Note. PRAGG = Peer-Nominated Aggression Scale; THAGG = Teachers' Predictions of Peer-Nominated Aggression Scale; CBCAGG = Child Behavior Checklist Aggression scale (as rated by teachers).
^a $n = 81$. ^b $n = 65$. ^c $n = 93$.
 * $p < .001$.

Table 3
Partial Correlations of Aggression Measures
Controlling for Grade ($N = 65$)

Measure	1	2	3
1. PRAGG	—		
2. THAGG	.66*	—	
3. CBCAGG	.49*	.49*	—

Note. PRAGG = Peer-Nominated Aggression Scale; THAGG = Teachers' Predictions of Peer-Nominated Aggression Scale; CBCAGG = Child Behavior Checklist Aggression scale (as rated by teachers).
 * $p < .001$.

late significantly with the comparable actual peer nominations, but these correlations are also all significantly lower than the correlations between the two measures of aggression. The teachers' predictions of popularity and victimization are particularly discrepant from the actual peer nominations.

Teachers may be particularly poor at predicting peer nominations on variables with stereotypic gender differences that are not recognized by the children themselves. A popular notion among teachers may be that girls are generally more prosocial and popular and less likely to be victimized, and this may influence their predictions of peer nominations. However, the children may not share this perception.

Support for this argument is provided by the mean scores for each gender, which are presented in Table 5. A multivariate analysis of variance was conducted with the six scales as dependent variables and with gender of child and type of rater (teacher vs. peer) as independent variables. A significant multivariate interaction was found between type of rater and gender of child, $F(6, 167) = 3.245, p < .005$, indicating that the gender differences in the teacher predictions of peer nominations were substantially different from the gender differences in the actual peer nominations. The significance of each pairwise difference as revealed by post hoc tests is shown in Table 5. As expected, boys

Table 4
Intercorrelations of Aggression Measures by
Format of Teacher Question

Measure	1	2	3
By question format			
1. PRAGG	—		
2. THAGG ^a	.83*	—	
3. CBCAGG ^b	.75*	.77*	—
By name format			
1. PRAGG	—		
2. THAGG ^c	.73*	—	
3. CBCAGG ^d	.61*	.56*	—

Note. PRAGG = Peer-Nominated Aggression Scale; THAGG = Teachers' Predictions of Peer-Nominated Aggression Scale; CBCAGG = Child Behavior Checklist Aggression scale (as rated by teachers).
^a $n = 55$. ^b $n = 53$. ^c $n = 119$. ^d $n = 77$.
 * $p < .001$.

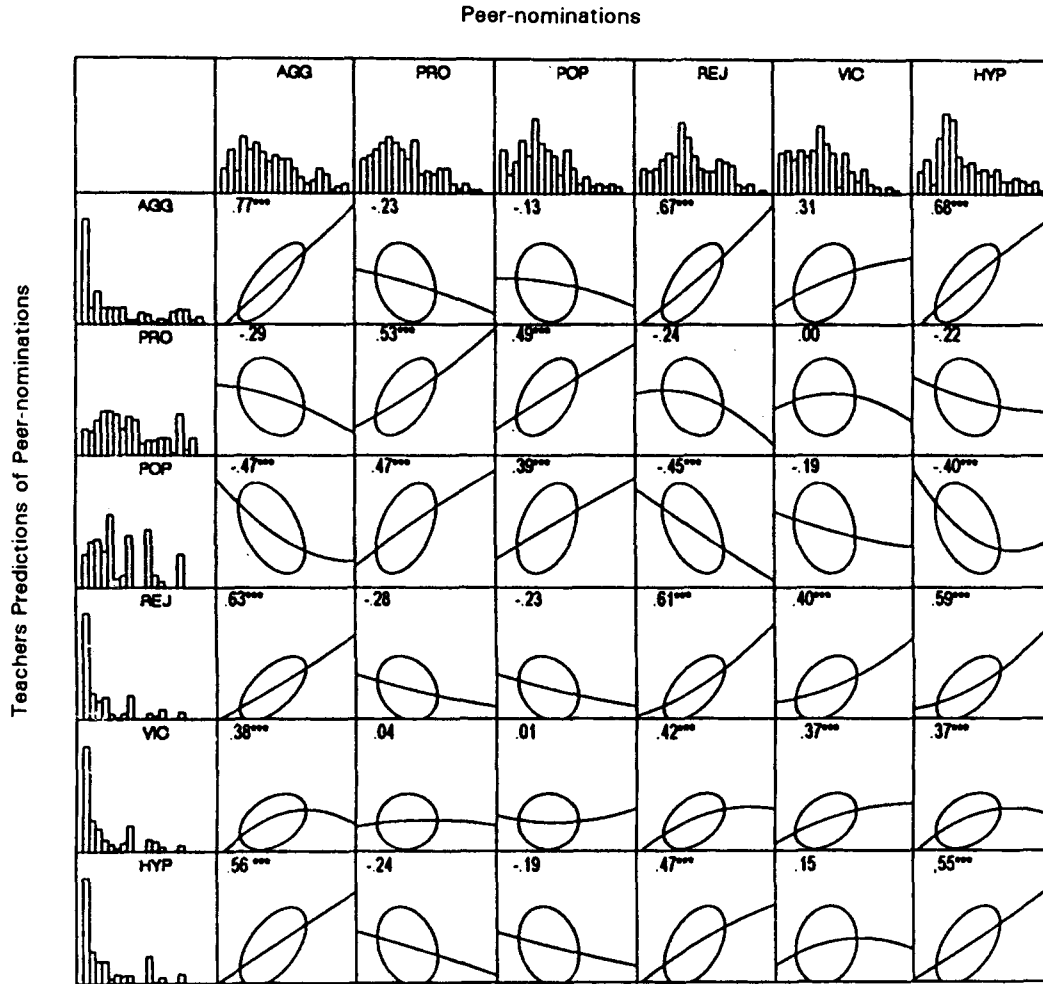


Figure 2. The correlations between teacher predictions of peer nominations and actual peer nominations on scales of Aggression (AGG), Prosocial Behavior (PRO), Popularity (POP), Rejection (REJ), Victimization (VIC), and Hyperactivity (HYP). ($N = 174$). The best-fitting quadratic curve and a 50% confidence ellipse are shown for each bivariate relation. *** $p < .001$.

and girls differed very significantly in their actual aggression and hyperactivity as assessed by peer nominations. Boys scored higher on both. The teachers' predictions of peer-nominated aggression and hyperactivity, although generally yielding lower mean scores, also show these same significant gender differences, providing construct validity for the teacher measure of aggression. However, the more important data for the argument that teachers are biased toward seeing girls as more prosocial, popular, and less victimized than boys are the mean gender differences on those three scales. As Table 5 reveals, teachers did predict that girls would be rated as significantly more prosocial and popular and less subject to victimization. However, the children did not nominate girls and boys significantly differently on any of these three variables; so, the explanation of teacher bias seems plausible.

Finally, in Table 6, the pattern of correlations between each aggression measure and the other behaviors measured the same way is displayed. One can see that the pattern of relations be-

tween teachers' predictions of aggressive behavior and other behaviors is remarkably similar to the pattern of relations for actual peer nominations. The one difference seems to be that the teachers overestimated the negative relation of aggressiveness to popularity and prosocial behavior and underestimated the positive relation between aggressiveness and rejection, victimization, and hyperactivity. Overall, this pattern of positive and negative correlations for the teacher measure of aggression confirms the differential validity of the teacher measure.

Discussion

A procedure was evaluated for using teachers to predict children's peer-nomination scores. Teachers were asked to predict the peer-nominated scores that children in their class would receive on aggression, hyperactivity, popularity, rejection, victimization, and prosocial behavior. About half of the teachers rated their children one at a time on all the questions (the by child

Table 5
Means and Significance Tests for Boys and Girls on Peer Nominations and Teachers' Predictions of Peer Nominations

Scale	Peer nominations			Teachers' predictions of peer nominations			Peer vs. teacher <i>F</i>	
	Boys (<i>n</i> = 81)	Girls (<i>n</i> = 93)	<i>F</i>	Boys (<i>n</i> = 81)	Girls (<i>n</i> = 93)	<i>F</i>	Boys	Girls
Aggression	.321	.225	14.68***	.197	.094	22.39***	95.61***	125.27***
Prosocial Behavior	.232	.263	1.56	.197	.244	4.26*	5.75*	1.28
Popularity	.173	.187	0.65	.168	.265	14.50***	0.12	16.75***
Rejection	.366	.298	5.94*	.123	.060	12.05***	197.57***	244.33***
Victimization	.209	.212	0.02	.110	.062	8.97**	33.20***	160.95***
Hyperactivity	.317	.260	4.47*	.151	.087	8.09**	80.77***	117.83***

* $p < .05$. ** $p < .01$. *** $p < .001$.

procedure); the other half rated all the children on one question, then on the next question, and so on (the by question procedure). The teachers also filled out the CBC for the children. In group sessions, without the teachers present, the children then completed peer nominations on the same items.

The results revealed that, in general, the teacher predictions were highly reliable and valid. The teachers were very good at predicting the peer-nominated aggression scores that children in their class would receive. They were less accurate at predicting peer nominations on other characteristics and particularly poor at predicting popularity and victimization. The teachers' predictions in general tended to be more skewed than actual peer nominations, with fewer children actually scoring at zero than the teachers predicted. The teacher predictions were as reliable (internally consistent) as the actual peer nominations.

The two versions of the procedure did not differ significantly in predictive validity, although there was a trend for the by question procedure to be better, as expected. The teacher predictions seemed to be just as valid for girls as for boys. Controlling for grade-related effects on aggression reduced the correlations with actual peer nominations 15% to 20%, but they still remained substantial.

The pattern of gender differences in aggression implied by the teachers' scores and the pattern of positive and negative relations of teachers' predictions of aggression with other teach-

ers' predictions were very similar to the relations observed among actual peer nominations and provide construct validity for the measure.

It is concluded that teachers' predictions of peer nominations of aggression can be used with confidence as a valid assessment measure to evaluate the aggression of elementary school children in settings in which actual peer-nomination procedures cannot be used. The teachers' predictions will not yield exactly the same scores as will actual peer nominations, but the individual differences reflected by the scores will be valid indicators of true individual differences in aggression. At the same time, teachers' predictions of nominations on other scales should be treated with caution. Teachers' predictions of popularity and victimization seem to be particularly problematic.

The teachers' predictions of peer nominations have some clear practical advantages for assessing aggression over using the CBC in many research settings. Most notably, high reliability and validity can be obtained with a 10-item scale out of 22 total questions in comparison with the 39-item scale and 120 total questions on the CBC. Although the norms for the checklist will often make it an instrument of choice for clinicians interested in identifying children with severe problems, the teachers' predictions of peer nominations can provide a score for each child that is more indicative of their behavioral problems as perceived by peers. Peer-nominated aggression in turn has been shown in many studies (Huesmann et al., 1984; Olweus, 1979) to be one of the very best early predictors of adult aggressive and antisocial behavior.

In our view, personality traits are best viewed as enduring behavioral predispositions. Aggressiveness is one such trait, and its measurement requires observation of behavior over time and across situations. For school children, peers are perhaps in the best position to engage in such observations; so it is not surprising that peer-nominated aggression has proved to be one of the most reliable and valid ways to assess childhood aggression. However, this article suggests that teachers' predictions of peer nominations run a close second. Teachers do observe children over time, although perhaps not in as varied situations as children. However, teachers are also privy to the peer group's knowledge about a child, and casting the teachers' task as predicting peer nominations allows them to use that information

Table 6
Correlations of Each Aggression Scale With Other Scales Assessed by the Same Technique

Other scales	Aggression scale	
	Peer nominations ^a	Teachers' predictions of peer nominations ^b
Prosocial Behavior	-.31**	-.44**
Popularity	-.22*	-.57**
Rejection	.91**	.79**
Victimization	.53**	.49**
Hyperactivity	.89**	.72**

^a $N = 176$. ^b $N = 174$.
* $p < .05$. ** $p < .001$.

as well. Consequently, teachers' predictions of peer nominations appear to be a valid and reliable alternative to using peer nominations for assessing the same enduring, aggressive dispositions in children that peer nominations detect.

References

- Achenbach, T. M., & Edelbrock, C. (1981). Behavioral problems and competencies reported by parents of normal and disturbed children aged four to sixteen. *Monographs of the Society for Research in Child Development*, 46 (Serial No. 188).
- Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, VT: University of Vermont Department of Psychiatry.
- Bjorkqvist, K., Lagerspetz, K., & Kaukiainen, A. (1992). Do girls manipulate and boys fight? Developmental trends in regard to direct and indirect aggression. *Aggressive Behavior*, 18, 117-127.
- Caprara, G. V. (1992, September). *Human aggression*. Paper presented at the meeting of the International Society for Research on Aggression, Siena, Italy.
- Dodge, K., & Coie, J. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology*, 53, 1146-1158.
- Eron, L. D., Laulicht, J. H., Walder, L. O., Farber, I. E., & Spiegel, J. P. (1961). Application of role and learning theories to the study of the development of aggression in children. *Psychological Reports*, 9, 291-334.
- Eron, L. D., Walder, L. O., & Lefkowitz, M. M. (1971). *Learning of aggression in children*. Boston: Little, Brown.
- Farrington, D. P. (1991). Childhood aggression and adult violence: Early precursors and later life outcomes. In D. J. Pepler & K. H. Rubin (Eds.), *The development and treatment of childhood aggression* (pp. 5-29). Hillsdale, NJ: Erlbaum.
- Guerra, N. G., Huesmann, L. R., Tolan, R. H., Van Acker, R., & Eron, L. D. (in press). Correlates of aggression among inner-city children: Implications for prevention. *Journal of Consulting and Clinical Psychology*.
- Huesmann, L. R., & Eron, L. D. (1986). *Television and the aggressive child: A cross-national comparison*. Hillsdale, NJ: Erlbaum.
- Huesmann, L. R., Eron, L. D., Lefkowitz, M. M., & Walder, L. O. (1984). The stability of aggression over time and generations. *Developmental Psychology*, 20, 1120-1134.
- Huesmann, L. R., Guerra, N. G., & Eron, L. D. (1992, November). *Victimization and aggression*. Paper presented at the meeting of the Society for Life History Research, Philadelphia, PA.
- Huesmann, L. R., Guerra, N. G., Eron, L. D., Zelli, A., Salamanca, K., Miller, L., & Hudson, G. (1989, November). *Changing children's cognitions about aggressive behavior*. Paper presented at the meeting of the American Society of Criminology, Reno, NV.
- Kazdin, A. E. (1987). Treatment of antisocial behavior in children: Current status and future directions. *Psychological Bulletin*, 102, 187-201.
- Lefkowitz, M. M., Eron, L. D., Walder, L. O., & Huesmann, L. R. (1977). *Growing up to be violent*. New York: Pergamon Press.
- Leibowitz, G. (1968). Comparison of self-report and behavioral techniques of assessing aggression. *Journal of Consulting and Clinical Psychology*, 32, 21-25.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 234-259.
- Offord, D. R., Boyle, M. C., & Racine, Y. A. (1991). The epidemiology of antisocial behavior in childhood and adolescence. In D. J. Pepler & K. H. Rubin (Eds.), *The development and treatment of childhood aggression* (pp. 31-54). Hillsdale, NJ: Erlbaum.
- Ollendick, T. H., Oswald, D. P., & Francis, G. (1989). Validity of teacher nominations in identifying aggressive, withdrawn, and popular children. *Journal of Clinical Child Psychology*, 18(3), 221-229.
- Olweus, D. (1979). The stability of aggressive reaction patterns in human males: A review. *Psychological Bulletin*, 85, 852-875.
- Perry, D. G., Perry, L. C., & Rasmussen, P. (1986). Cognitive social learning mediators of aggression. *Child Development*, 57, 700-711.
- Pulkkinen, L. (1987). Offensive and defensive aggression in humans: A longitudinal perspective. *Aggressive Behavior*, 13, 197-212.
- Walder, L. O., Abelson, R. P., Eron, L. D., Banta, T. J., & Laulicht, J. H. (1961). Development of a peer-rating measure of aggression. *Psychological Reports*, 9, 497-556.
- Wiggins, J. S., & Winder, C. L. (1961). The Peer Nomination Inventory: An empirically derived sociometric measure of adjustment in preadolescent boys. *Psychological Reports*, 9, 643-677.
- Williams, J. R., Meyerson, L. J., Eron, L. D., & Semler, S. J. (1967). Peer-rated aggression and aggressive responses elicited in an experimental situation. *Child Development*, 38, 181-190.

Received November 23, 1992

Revision received March 11, 1994

Accepted March 28, 1994 ■