

The Aggression Scale: A Self-Report Measure of Aggressive Behavior for Young Adolescents

Pamela Orpinas
University of Georgia
Ralph Frankowski
University of Texas–Houston

This article describes the development and psychometric properties of the Aggression Scale. The scale consists of 11 items designed to measure self-reported aggressive behaviors among middle school students (sixth, seventh, and eighth graders). The scale was evaluated in two independent samples of young adolescents (n = 253 and n = 8,695). Reliability scores were high in both samples, and did not vary significantly by gender, ethnicity, or grade level in school. Aggression scores also were stable in a 2-year follow-up study. Mean scores on the Aggression Scale were associated positively with teachers' independent rating of student aggression, other measures of aggression, and known predictors of aggression. The scale is brief, is easy to administer, and focuses on overt behaviors. Thus, the Aggression Scale could be a useful tool for program evaluation and for further research on violence prevention in schools.

Violence is one of the most prevalent and destructive behaviors that adolescents face today, because they are at particular risk of being either the victim or the perpetrator of an act of violence (Maguire & Pastore, 1998). Schools have an especially important role in violence prevention, because a large proportion of youth can be reached through schools, and violence is a problem that affects schools directly (American School Health Association, Association for the Advancement of Health Education, Society for Public Health Education, 1989; Callahan & Rivara, 1992). In addition, early intervention with students who display aggressive behavior is important because they are

This research was partially supported by the Centers for Disease Control and Prevention, National Center for Injury Prevention and Control (Grant: U81/CCU609953-02) and by the Hogg Foundation for Mental Health (Grant: 2901H).

Journal of Early Adolescence, Vol. 21 No. 1, February 2001 50-67

© 2001 Sage Publications, Inc.

at substantial risk for future violent behavior, delinquency, and school withdrawal (Eron, 1987; Kupersmidt & Coie, 1990).

To curb violence among students, schools have been implementing a variety of violence prevention programs and curricula. One of the problems in evaluating those programs is the lack of appropriate instruments for young adolescents. Instruments have been developed to evaluate aggression among special populations, such as psychiatric in-patients (Kazdin, Rodgers, Colbus, & Siegel, 1987; Korn et al., 1992; Morrison, 1993; Plutchik & van Praag, 1990), incarcerated boys (Oyserman & Saltz, 1993), or children with attention deficit disorder (Moffitt & Silva, 1988). Some instruments have been developed to evaluate criminal or antisocial behavior (Tremblay, Pihl, Vitaro, & Dobkin, 1994; Weis, 1986), rather than nondelinquent aggressive behavior. Scales based on teacher ratings (Matthews & Angulo, 1980), peer ratings (Lardon & Jason, 1992), or family observations (Yudofsky, Silver, Jackson, Endicott, & Williams, 1986) are more expensive and less practical in large prevention projects than are self-reports. Some instruments have been developed for college students (Buss & Durkee, 1957; Buss & Perry, 1992) or for high school students (Kolbe, Kann, & Collins, 1993). The few scales that measure self-reported aggression during early adolescence, however, do not measure the frequency of self-reported behavior. They are designed to evaluate students' perceptions as to the extent to which a specific behavior applies to themselves (Achenbach et al., 1990; Cotten et al., 1994) or they are designed to evaluate behavioral intention (Deluty, 1979). Thus, to the authors' knowledge, no scale is available to measure specifically the frequency of self-reported aggressive behaviors among middle school students. Measuring self-reported behaviors is important in violence prevention research because knowledge, attitudes, beliefs, and behavioral intentions frequently have an unknown or tenuous association with related risk behaviors (Kolbe et al., 1993). In addition, even though changing attitudes or knowledge might be important, the final desired outcome of any violence prevention program is to reduce the frequency of aggressive behaviors.

The purpose for this article is to describe the development and the psychometric properties of the Aggression Scale, which during a specific time frame measures frequency of self-reported aggressive behaviors among middle school students. Three studies are reported. In Study 1, the development of the scale, its content validity, and the consistency of students' responses are presented. In Study 2, the construct validity and internal consistency of the scores based on a small sample of sixth graders is described. In Study 3, the properties of the scale were cross-validated in a large sample of sixth, seventh, and eighth graders. Reliability was evaluated using test-retest as a measure of stability over time (Study 3) and Cronbach's alpha as a measure of internal con-

sistency (Studies 2 and 3). Content validity was assessed by experts in the field (Study 1). Construct validity (Studies 2 and 3) was measured as the relation of the scale's scores with an independent measure of aggression and with other measures consistent with theoretically derived hypotheses (Carmines & Zeller, 1989). The hypotheses evaluated were that the scores of the Aggression Scale would be associated positively with an independent teacher rating of aggressive behaviors and with other items that measure aggression, such as physical fights at school, injuries due to fights, and weapon-carrying. Based on the literature, it was expected that the aggression scores would be associated positively with alcohol and substance abuse (Orpinas, Basen-Engquist, Grunbaum, & Parcel, 1995; Valois, Vincent, McKeown, Garrison, & Kirby, 1993), low academic achievement (Farrington, 1989), and low parental monitoring (Orpinas, Murray, & Kelder, 1999; Patterson & Stouthamer-Loeber, 1984).

STUDY 1: SCALE DEVELOPMENT

Generation of Items

To develop the Aggression Scale, a list of common verbal and physical aggressive behaviors among middle school students was created. To ensure that the items reflected the most common aggressive acts among students and that the wording of the items was appropriate, items were reviewed by 1 school counselor, 12 teachers (3 men and 9 women from six middle schools), and 2 university professors with expertise in school health promotion and violence prevention. In addition to physical and verbal aggression, consultants suggested two additional items to measure anger.

The original scale of 17 items was administered to 85 students (46 boys, 39 girls; mean age = 11.8 years) in three sixth-grade classes at two schools. The student survey, as well as the use of passive informed consent, was approved by the Committee for the Protection of Human Subjects of the University of Texas–Houston and by the school district's research department. Eight parents returned the consent form requesting their child not participate, and one student refused to participate. In two of the three classes ($n = 56$), the 17-item scale was administered twice to each student with a 1-week interval between administrations. After the first administration, students gave feedback regarding the content and format of the items. The 17-item scale was analyzed for clarity of instructions, organization of the items, and amount of time required for administration. Items were analyzed for ambiguity and com-

prehension. During this pilot evaluation, 6 items were deleted from the original 17-item Aggression Scale. Three items were deleted because of their low stability over time; 1 item was deleted because it did not measure interpersonal aggression; 1 item was deleted because it did not add new information and had a double meaning; and 1 item was deleted because students rated the behavior as not safe nor necessarily aggressive.

In the third class ($n = 29$), the final form of the scale (11 items) was administered twice with a 1-day interval to measure the temporal consistency of students' responses; 23 students were present in both administrations. This short time frame was used because the scale requests information about aggressive behaviors during the week prior to the administration of the survey. The consistency of responses, evaluated by a paired t test to compare the mean difference between the two evaluations, was high (mean difference = 0.0; 95% confidence interval [CI] = 2.49, -2.49; $n = 23$). The intraclass correlation between the two evaluations, which is an estimate of reliability, also was high (intraclass correlation = .85). All analyses were done using SPSS/PC.

Final Scale and Scoring Procedures

The final Aggression Scale consists of 11 items (see appendix). The scale measures behaviors that might result in psychological or physical injury to other students. Although the instructions do not specify the setting and, therefore, the behaviors can occur in or out of school, most questions refer to aggressive behaviors against other students. The scale is not intended to measure other forms of aggression such as family violence, aggression against teachers, or destruction of property. Questions are limited to overt aggressive behaviors and do not include relational aggression, that is, behaviors that harm others through peer relationships (excluding others from a playgroup or withdrawing friendship [e.g., Crick, 1995]). The scale requests information regarding the frequency of the most common overt aggressive behaviors, including verbal aggression (teasing, name-calling, encouraging students to fight, threatening to hurt or hit) and physical aggression (pushing, slapping, kicking, hitting), as well as information about anger (getting angry easily, being angry most of the day). These components were summated into a single scale. The measurement of such aggressive behaviors is fundamental in school violence prevention research. Those behaviors not only are unacceptable in schools but also might be indicators of low self-control, which is one of the personal characteristics associated with delinquency (Gottfredson & Hirschi, 1990). To minimize recall bias, the scale requests information about behaviors during the past 7 days. Responses to each item can range from 0

times through 6 or more times. Responses are additive; thus, the Aggression Scale ranges between 0 and 66 points. The instructions for completing the scale are given orally by the person administering the scale.

STUDY 2: VALIDITY AND INTERNAL CONSISTENCY

Method

Participants. Data were collected from 10 sixth-grade classes of four middle schools of a large metropolitan public school district in Texas. Of the 265 students in these classes, 253 (115 boys, 138 girls) completed the Aggression Scale. This sample was the baseline evaluation of a study designed to evaluate a violence-prevention program (Orpinas, Parcel, McAlister, & Frankowski, 1995). Of the respondents, 63% were Hispanic ($n = 160$), 17% were African American ($n = 44$), and 18% were Caucasian ($n = 45$); 90% of the students were between 11 and 12 years of age. One half of the students surveyed were eligible for free or reduced-price lunch; a similar proportion of students was eligible for free or reduced-price lunch in the district's total student population.

Measures. The survey that was administered to participating students was composed of the Aggression Scale, other measures of aggression, and predictors of aggression. Three questions taken from the Centers for Disease Control and Prevention's Youth Risk Behavior Survey (Kolbe et al., 1993) measured injuries due to fights, weapon-carrying, and alcohol use. The time frame for the three questions was the month prior to the survey. Frequency of injuries due to fights ranged between 0 and 4 or more times, frequency of weapon-carrying ranged between 0 and 6 or more days, and frequency of alcohol use ranged between 0 and 30 days.

Parental monitoring was measured by two questions: "Do your parents let you come and go as you please?" and "When you are away from home, do your parents know where you are and who you are with?" Students had five possible responses, ranging between *never/almost never* and *almost always*. Responses to both questions were added and recoded into five categories ranging between *very high* and *very low* parental monitoring.

Academic achievement was reported by the teachers and was measured as the student's most usual grades. Teachers graded students as obtaining *mostly As and Bs, mostly Bs and Cs, mostly Cs and Ds, or mostly Ds and Fs*.

Procedures. The student survey, as well as the use of passive informed consent, was approved by the Committee for the Protection of Human Subjects of the University of Texas–Houston and by the school district’s research department. Before the administration of the questionnaire, a parental consent form was sent to parents explaining the purpose and content of the evaluation. Parents who did not wish that their child participate in this evaluation could sign the letter and return it to the school. Research staff administered the questionnaire following standardized instructions. Students were informed about the importance of the survey and that completing it was voluntary.

Results

Construct validity. The relation between self-reported aggression, measured by the Aggression Scale, and other self-reported measures and predictors of aggression were analyzed. Univariate analysis of variance showed a positive relation between the mean score in the Aggression Scale and the number of injuries due to fights, $F(3, 245) = 14.1$, and the number of days students carried a weapon, $F(3, 248) = 16.1$. Mean aggression scores also increased as the number of days students reported drinking alcohol increased, $F(3, 248) = 16.1$, parental monitoring decreased, $F(4, 244) = 6.0$, and academic achievement decreased, $F(3, 244) = 7.1$ (see Table 1). All these relations were statistically significant ($p < .0001$). The mean score on the Aggression Scale was significantly higher among boys ($\bar{X} = 19.3$, $SD = 15.5$) than among girls ($\bar{X} = 13.2$, $SD = 12.9$), $F(1, 251) = 11.7$, $p = .0007$.

Internal consistency. The internal consistency scores, estimated with Cronbach’s alpha coefficient, were high (.87 for the total sample, .88 for boys, .87 for girls, .86 for Hispanic students, .85 for African American students, and .92 for Caucasian students). The item-deleted estimates for alpha coefficients for the total sample ranged from .85 through .87.

STUDY 3: CROSS-VALIDATION

Method

Participants and procedures. In 1994, data were collected from all sixth, seventh, and eighth graders of eight middle schools of a large metropolitan public school district in Texas; 8,695 students (87% response rate) completed

TABLE 1: Mean Scores on the Aggression Scale for Selected Variables

Selected Variable	Study 2			Study 3								
	Sixth Graders			Sixth Graders			Seventh Graders			Eighth Graders		
	n	\bar{X}	(SD)	n	\bar{X}	(SD)	n	\bar{X}	(SD)	n	\bar{X}	(SD)
Physical fights at school												
0 times	—	—	—	2,369	12.7	(12.4)	2,147	13.3	(12.2)	2,185	14.4	(13.2)
1 time	—	—	—	446	20.7	(14.6)	281	21.4	(13.4)	259	20.2	(14.1)
2 to 3 times	—	—	—	200	28.5	(15.0)	180	28.2	(15.4)	149	29.5	(14.2)
4 or more times	—	—	—	169	32.2	(17.0)	126	32.4	(16.8)	144	34.9	(18.2)
Injuries due to fights												
0 times	226	14.5	(13.1)	2,701	14.5	(13.6)	2,345	14.7	(13.3)	2,394	15.5	(13.9)
1 time	12	21.1	(17.4)	277	20.8	(15.3)	202	21.2	(13.9)	168	21.9	(15.6)
2 to 3 times	7	35.0	(16.6)	130	24.9	(16.8)	118	24.6	(14.9)	110	27.7	(16.0)
4 or more times	4	48.5	(10.9)	84	28.2	(17.2)	73	29.2	(17.3)	66	36.8	(17.2)
Weapon-carrying												
0 days	218	13.9	(13.2)	2,280	12.6	(12.3)	1,900	12.5	(11.8)	1,878	12.5	(11.9)
1 day	16	20.9	(14.6)	230	17.6	(12.6)	197	19.3	(13.4)	180	19.7	(12.7)
2 to 3 days	8	37.0	(11.7)	221	24.0	(14.9)	220	20.7	(12.8)	204	22.1	(13.3)
4 or more days	10	35.1	(14.7)	353	29.4	(16.9)	358	29.4	(16.1)	458	31.2	(16.4)
Alcohol use												
0 days	212	14.2	(13.5)	2,154	12.9	(12.5)	1,783	12.8	(12.1)	1,587	12.5	(12.2)
1 to 2 days	30	22.7	(14.4)	330	22.5	(14.2)	379	19.9	(13.4)	488	18.5	(13.5)
3 to 9 days	7	37.1	(15.5)	160	26.2	(15.7)	222	26.8	(14.9)	326	24.6	(14.4)
10 or more days	3	22.0	(19.5)	100	32.5	(17.4)	101	32.7	(17.3)	207	33.1	(17.2)

Marijuana use												
0 times	—	—	—	2,344	13.8	(12.8)	1,971	13.4	(12.2)	1,850	13.2	(12.4)
1 to 2 times	—	—	—	159	26.2	(16.1)	199	22.4	(14.3)	243	21.3	(13.8)
3 to 9 times	—	—	—	54	27.1	(14.4)	108	27.9	(14.8)	159	23.1	(14.9)
10 or more times	—	—	—	115	31.8	(18.0)	173	31.2	(15.9)	325	30.6	(16.7)
Parental monitoring												
Very high	35	10.5	(10.0)	550	10.4	(11.1)	384	11.4	(11.9)	363	10.3	(10.8)
High	115	15.1	(13.6)	1,349	14.0	(12.6)	1,148	13.1	(11.9)	1,049	14.5	(13.0)
Medium	77	17.2	(15.3)	938	18.2	(15.3)	875	18.2	(14.2)	972	18.7	(15.3)
Low	17	23.4	(15.0)	276	24.1	(16.9)	257	25.1	(16.6)	296	25.4	(16.0)
Very low	5	38.0	(16.4)	50	34.0	(16.2)	58	27.5	(15.7)	47	31.7	(21.1)
Grades												
Mostly As and Bs	120	13.5	(13.4)	1,383	13.3	(12.9)	976	12.3	(12.3)	928	13.1	(13.3)
Mostly Bs and Cs	84	14.7	(13.4)	1,325	16.1	(14.2)	1,264	16.8	(13.6)	1,302	17.2	(14.4)
Mostly Cs and Ds	30	22.2	(14.7)	388	21.3	(16.4)	383	20.4	(15.4)	408	21.9	(15.2)
Mostly Ds and Fs	14	28.1	(17.4)	96	27.3	(18.2)	98	24.4	(17.1)	92	27.7	(19.4)

NOTE: Dashes indicate that data were not collected. *p* values were < .0001 for all comparisons.

the Aggression Scale. None of these students participated in Study 2. This sample was a baseline evaluation of a large study designed to evaluate a violence prevention intervention (Kelder et al., 1996; Orpinas et al., 2000). No formal violence prevention program had been introduced at the time of the evaluation. The sample was ethnically diverse: 66% Hispanic ($n = 5,707$), 19% African American ($n = 1,676$), 8% Caucasian ($n = 717$), 4% Asian ($n = 305$), and less than 1% Native American ($n = 61$). The mean age of students was 13 years (12 years in sixth grade, 13 years in seventh grade, and 14 years in eighth grade).

Students who were in sixth grade in 1994 were followed through seventh grade in 1995 and eighth grade in 1996. Stability over time was calculated in a subsample of those students ($n = 885$; 435 boys, 450 girls). The subsample consisted of students who completed all three evaluations and who attended schools without a formal violence prevention program. The ethnic distribution of this subsample was similar to the 1994 sample.

To obtain an independent measure of students' aggressive behavior, teachers of eighth-grade classes of four schools were asked to rate students' aggressive behavior. Three classes per school were selected randomly for that purpose. Three teachers at each school were asked to rate independently each student's behavior. Teachers of differing subject areas were asked to participate. Teachers ranked students in a 4-point scale: 0 = *not aggressive*; 1 = *a little aggressive, once or twice a week is aggressive toward other students*; 2 = *somewhat aggressive, several times a week is aggressive toward other students or has some difficulty controlling his or her anger*; and 3 = *aggressive, frequently is aggressive toward other students or is usually very angry*. Physical and verbal aggression, as well as anger, were defined using the same behaviors described in the Aggression Scale. Teachers received written instructions on how to rate students. Teachers evaluated 255 eighth-grade students (113 boys, 133 girls, and 9 students who did not identify their gender). Those 255 students completed a survey containing the Aggression Scale.

The student survey and the teacher evaluation, as well as the use of passive informed consent, was approved by the Committee for the Protection of Human Subjects of the University of Texas–Houston and by the school district's research department. The student data collection process was similar to the process described in Study 2.

Measures. The survey that was administered to participating students was composed of the Aggression Scale, other measures of aggression, and predictors of aggression. Five questions taken from the Centers for Disease Control

and Prevention's Youth Risk Behavior Survey measured fights at school (prior month), injuries due to fights (prior year), weapon-carrying (prior month), alcohol use (prior month), and marijuana use (lifetime). Weapon-carrying was measured by four questions. Students were asked how many times during the past 30 days they have carried: (a) a handgun, (b) other guns (such as a rifle or shotgun), (c) a knife or razor, and (d) a club, stick, or pipe. Students' responses were combined to calculate one overall estimate of weapon-carrying. A study conducted by Brener, Collins, Kahn, Warren, and Williams (1995) with a sample of adolescents showed that test-retest reliability scores were substantial for lifetime use of marijuana ($\kappa = .88$), weapon-carrying ($\kappa = .76$), alcohol use ($\kappa = .68$), and physical fights ($\kappa = .68$), and somewhat lower for injuries ($\kappa = .51$) (Brener et al., 1995).

The evaluation of parental monitoring followed the same procedure described in Study 2. To evaluate academic achievement, students' self-reported their most usual grades. Students indicated whether their grades were *mostly As and Bs*, *mostly Bs and Cs*, *mostly Cs and Ds*, or *mostly Ds and Fs*.

Results

Construct validity. A mean score for the three teacher ratings was calculated for each student and compared with the students' scores on the Aggression Scale. Mean scores on the Aggression Scale significantly increased as teacher ratings increased, $F(6, 222) = 4.15, p < .001$ (effect size = .30). When the mean for the teacher rating was 0, the students' mean score on the Aggression Scale was 11.7; when the mean for the teacher rating was 1, the students' mean score on the Aggression Scale was 19.2; and when the mean for the teacher rating was 2 or more, the students' mean score on the Aggression Scale was 27.8. Both for boys and girls, the students' mean score on the Aggression Scale increased as the mean for the teacher rating increased.

Univariate analysis of variance showed a positive relation between the mean score on the Aggression Scale and self-reported number of fights, $F(3, 8656) = 535.0, p < .0001$; number of injuries due to fights, $F(3, 8669) = 189.4, p < .0001$; and number of days students carried a weapon, $F(3, 8481) = 668.6, p < .0001$. The relation between self-reported aggression and four strong predictors of aggression—alcohol use, marijuana use, low parental monitoring, and low academic achievement—also was analyzed. As expected, results from analysis of variance and test for linear trends showed that mean score on the Aggression Scale increased as self-reported alcohol consumption increased, $F(3, 7839) = 505.4, p < .0001$; marijuana use increased, $F(3, 7702) =$

451.1, $p < .0001$; parental monitoring decreased, $F(4, 8612) = 223.2$, $p < .0001$; and academic achievement decreased, $F(3, 8644) = 163.37$, $p < .0001$ (see Table 1). The mean score on the Aggression Scale was higher significantly among boys ($\bar{X} = 18.0$, $SD = 15.1$) than among girls ($\bar{X} = 14.4$, $SD = 13.4$), $F(1, 8,688) = 140.3$, $p < .0001$.

Exploratory and confirmatory factor analyses. To examine the internal structure of the Aggression Scale, the total sample was divided randomly in half. The first half ($n = 4,337$) was used for the exploratory factor analysis. Data were analyzed using SPSS. The 11-item scale was examined using exploratory factor analysis with maximum likelihood extraction followed by oblique rotation and Kaiser normalization. Due to the high correlation between items, oblique rotation was used to evaluate the empirical relation among the components of the Aggression Scale. Two factors comprising all 11 items were retained (eigenvalues > 1). The first factor, named Physical and Verbal Aggression, had 9 items and accounted for 40.6% of the total variance. The second factor, named Anger, had 2 items and accounted for an additional 5.0% of the total variance. One item, "I got into a physical fight because I was angry," had a strong loading on both factors (see Table 2). Pearson correlation between the two factors was .50; the correlation between Anger and the total Aggression Scale was .67; and the correlation between Physical and Verbal Aggression and the total Aggression Scale was .98.

The second half of the sample ($n = 4,446$) was used for the confirmatory factor analysis. Data were analyzed using LISREL 8.2 (Jöreskog & Sörbom, 1993, 1996). The model included the two factors identified in the exploratory factor analysis. The standardized factor loadings were comparable among items, ranging between .54 and .74 for Physical and Verbal Aggression and between .63 and .70 for Anger. The correlation between the two factors was high (.70). The minimum fit function chi-square test was significant (962.5; $p < .001$), which was expected given that this study was based on a very large sample size and that the chi-square test is highly dependent on sample size. Other indicators, however, did indicate a good fit of the model. The root-mean-square residuals (RMR) was low (.037) and the goodness-of-fit index (GFI) was high (GFI = .96 and adjusted GFI = .94), both indicating a good overall fit of the model. Values for the RMR below .05 indicate a good fit. The GFI ranges between 0 (no fit) and 1 (perfect fit), and values above .90 indicate a good fit (Byrne, 1989). A specific area of misspecification of the model was related to the item "I got into a physical fight because I was angry," which had the largest modification index and was part of a number of large standardized residuals.

TABLE 2: Structure Coefficients of the Aggression Scale: Results of the Exploratory Factor Analysis

<i>Items Grouped Into Factors</i>	<i>Structure Coefficients</i>	
	<i>Factor 1</i>	<i>Factor 2</i>
Physical and Verbal Aggression		
I teased students to make them angry	.66	.35
I fought back when someone hit me first	.60	.44
I said things about other kids to make other students laugh	.69	.34
I encouraged other students to fight	.62	.40
I pushed or shoved other students	.76	.44
I slapped or kicked someone	.68	.49
I called other students bad names	.73	.39
I threatened to hurt or to hit someone	.72	.48
I got into a physical fight because I was angry	.51	.47
Anger		
I got angry very easily with someone	.46	.64
I was angry most of the day	.40	.72

NOTE: Bold numbers indicate that those items have a strong factor loading on that factor.

Internal consistency. The internal consistency of the scores, estimated with Cronbach's alpha coefficient, was high. Cronbach's alpha was .88 for the total sample. The analysis was repeated for boys and for girls separately, for each ethnic group, and for each grade level in school. A similar Cronbach's alpha coefficient was found for each subsample, ranging between .86 and .88.

Stability over time. Stability over time was evaluated by a paired *t* test comparing the mean difference between pairs of evaluations with a 1-year and a 2-year difference. Mean differences were not significant statistically (1994-1995: mean difference = -0.21, 95% CI = -0.94, 0.52, *n* = 874; 1994-1996: mean difference = 0.05, 95% CI = -0.81, 0.92, *n* = 876; 1995-1996: mean difference = 0.28, 95% CI = -0.52, 1.08, *n* = 882). Correlation coefficients between pairs of evaluations were fairly high and, as expected, were higher over the 1-year follow-up than over the 2-year follow-up (1994-1995: *r* = .63; 1995-1996: *r* = .56; 1994-1996: *r* = .50).

DISCUSSION

The Aggression Scale was developed to evaluate self-reported aggressive behavior among middle school students. Described in three studies reported

here are the development and psychometric properties of the scale in diverse samples of middle school students. Across samples, scores showed high internal consistency and stability over time and were consistent with an external teacher evaluation of aggression. However, these results must be viewed in the light of some limitations on the generality of the results. All participants were middle school students of a large urban school district in Texas and, thus, results might not be generalizable to other populations or to other regions of the United States. The authors currently are evaluating the psychometric properties of the scale among a population of students living in rural areas. The application of the scale to other contexts besides schools remains an area for further research.

Another possible limitation is the use of self-report. The controversy on whether the validity of self-report indices is better or worse than non-self-report indices is still alive. However, several studies have shown that the construct validity of self-reports is superior to the validity of other measurement approaches (Howard, 1994), and self-reports have the great advantage of easy and inexpensive data collection.

In the development of the Aggression Scale, items that represent three components were included. The first two components, Physical Aggression and Verbal Aggression, are direct acts of aggression toward others. The third component, Anger, represents the emotional arousal that can lead to aggression (Berkowitz & Heimer, 1989). These three components have been included also in other scales (Buss & Perry, 1992). As revealed by the exploratory factor analysis, physical and verbal aggression accounted for 40.6% of the variance, with the two items related to anger comprising an additional 5.0%. The confirmatory factor analysis showed that the standardized factor loadings were comparable among items and that both factors were correlated highly; thus, it supports having a single scale encompassing both components. The strong association between the factors and the total aggression scores, as well as the high internal consistency of the scores, gives additional support to using the scale as a measure of a single construct.

The internal consistency scores of the Aggression Scale were high and did not vary significantly by gender, ethnicity, or grade level in school, indicating practical equivalency of scores among differing subgroups. The advantage of this high internal consistency is that researchers will need a smaller sample size to detect meaningful differences (DeVellis, 1991). Although the absolute prevalence of aggressive behaviors did vary by gender and ethnicity (Orpinas et al., 1999), the psychometric characteristics of the scale did not change.

The stability of the construct over 1 and over 2 years also was high. A similar stability of aggressive and delinquent behaviors has been found by other researchers (e.g., Olweus, 1979; Tremblay et al., 1994). Although the mean

difference in aggressive behaviors was not statistically different over the 2-year period, individual differences were observed. Not all aggressive children continue on the path of aggression, and some children will start a path of aggression at later years. These individual differences were evident from the correlation coefficients, which showed a strong but not perfect association. Correlation coefficients might be high, although the absolute prevalence, seriousness, or frequency can change over time. The scale could be a useful tool to identify which students desist from, or engage in, aggression as they grow older. The scale, however, does not provide psychological information about why students' aggression might increase or decrease. The scale provides a measure to monitor aggressive behaviors in schools, independent of whether the behavior is an expression of a serious psychological disorder, of a passing state, or of a learned behavior that has been rewarded by the student's environment.

Content validity was analyzed at three levels: experts from the university, teachers and counselors with experience working with students, and by the students themselves. They all contributed to the development of the items from their individual perspectives. Construct validity was measured by the association between the aggression scores and teachers' evaluations of students' aggressive behaviors, responses to other violence-related items, and responses to items considered in the literature as strong predictors of violence. The association between students' mean aggression scores and teachers' ratings was strong: Aggression scores almost tripled as the mean rating of teachers increased. As expected, mean aggression scores significantly and substantially increased with the number of fights at school, the number of days students carried a weapon, and the number of injuries due to fights. These three variables have been used by the Centers for Disease Control and Prevention to monitor youth violence (Kolbe et al., 1993).

The association between the aggression scores and other variables described in the literature as predictors of aggression provided additional information about the validity of the scores. As expected, mean aggression scores increased with the self-reported number of days students drank alcohol, the number of times students smoked marijuana, the perception of low parental monitoring, and lower academic achievement. For example, the mean aggression score was 2.6 times higher among students who reported drinking alcohol 10 or more days per month than among students who did not report alcohol consumption. The increase in aggression scores for all the variables not only was significant statistically but also substantially important. An increase in aggression scores from 12 (that is, 12 aggressive acts per week) to 30 aggressive acts per week is of real practical significance; it is the difference between an already difficult school and a nightmare.

To summarize, the scores of the Aggression Scale in these diverse samples of middle school students were highly stable over time, internally consistent, and associated positively with teacher evaluations. The Aggression Scale is practical in that it takes only a few minutes to complete and is well understood by students. The scale also focuses on overt behaviors, which are easier to quantify than subjective perceptions. Overt behaviors also should be the focus of violence prevention efforts. Thus, the Aggression Scale could be a useful tool for program evaluation and further research on violence prevention in middle schools.

APPENDIX The Aggression Scale

Please answer the following questions thinking of what you actually did during the last 7 days. For each question, mark with a circle how many times you did that behavior during the last 7 days.

<i>During the last 7 days</i>	<i>0 times</i>	<i>1 time</i>	<i>2 times</i>	<i>3 times</i>	<i>4 times</i>	<i>5 times</i>	<i>6 or more times</i>
1. I teased students to make them angry.	0	1	2	3	4	5	6+
2. I got angry very easily with someone.	0	1	2	3	4	5	6+
3. I fought back when someone hit me first.	0	1	2	3	4	5	6+
4. I said things about other kids to make other students laugh.	0	1	2	3	4	5	6+
5. I encouraged other students to fight.	0	1	2	3	4	5	6+
6. I pushed or shoved other students.	0	1	2	3	4	5	6+
7. I was angry most of the day.	0	1	2	3	4	5	6+
8. I got into a physical fight because I was angry.	0	1	2	3	4	5	6+
9. I slapped or kicked someone.	0	1	2	3	4	5	6+
10. I called other students bad names.	0	1	2	3	4	5	6+
11. I threatened to hurt or to hit someone.	0	1	2	3	4	5	6+

REFERENCES

- Achenbach, T. M., Bird, H. R., Canino, G., Phares, V., Gould, M. S., & Rubio-Stipec, M. (1990). Epidemiological comparisons of Puerto Rican and US mainland children: Parent, teacher, and self-reports. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*, 84-93.
- American School Health Association, Association for the Advancement of Health Education, Society for Public Health Education. (1989). *The National Adolescent Student Health Survey: A report on health of America's youth*. Oakland, CA: Third Party Publishing.
- Berkowitz, L., & Heimer, K. (1989). On the construction of the anger experience: Aversive events and negative priming in the formation of feelings. *Advances in Experimental Social Psychology, 22*, 1-37.
- Brener, N. D., Collins, J. L., Kahn, L., Warren, C. W., & Williams, B. I. (1995). Reliability of the Youth Risk Behavior Survey questionnaire. *American Journal of Epidemiology, 141*, 575-580.
- Buss, A. H., & Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology, 21*, 343-349.
- Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology, 63*, 452-459.
- Byrne, B. M. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. New York: Springer Verlag.
- Callahan, C. M., & Rivara, F. P. (1992). Urban high school youth and handguns. *Journal of the American Medical Association, 267*, 3038-3042.
- Carmines, E. G., & Zeller, R. A. (1989). *Reliability and validity assessment*. Newbury Park, CA: Sage.
- Cotten, N. U., Resnick, J., Browne, D. C., Martin, S. L., McCarragher, D. R., & Woods, J. (1994). Aggression and fighting behavior among African-American adolescents: Individual and family factors. *American Journal of Public Health, 84*, 618-622.
- Crick, N. R. (1995). Relational aggression: The role of intent attributions, feelings of distress, and provocation type. *Development and Psychopathology, 7*, 313-322.
- Deluty, R. H. (1979). Children's action tendency scale: A self-report measure of aggressiveness, assertiveness, and submissiveness in children. *Journal of Consulting and Clinical Psychology, 47*, 1061-1071.
- DeVellis, R. F. (1991). *Scale development: Theory and applications*. Newbury Park, CA: Sage.
- Eron, L. D. (1987). The development of aggressive behavior from the perspective of developing behaviorism. *American Psychologist, 42*, 435-442.
- Farrington, D. P. (1989). Early predictors of adolescent aggression and adult violence. *Violence and Victims, 4*, 79-100.
- Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford, CA: Stanford University Press.
- Howard, S. H. (1994). Why do people say nasty things about self-reports? *Journal of Organizational Behavior, 15*, 399-404.
- Jöreskog, K. G., & Sörbom, D. (1993). *Structural equation modeling with SIMPLIS command language*. Chicago: Scientific Software International.
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago: Scientific Software International.
- Kazdin, A. E., Rodgers, A., Colbus, D., & Siegel, T. (1987). Children's hostility inventory: Measurement of aggression and hostility in psychiatric in-patient children. *Journal of Clinical Child Psychology, 16*, 320-328.

- Kelder, S. H., Orpinas, P., McAlister, A., Frankowski, R., Parcel, G. S., & Friday, J. (1996). The Students for Peace Project: A comprehensive violence-prevention program for middle school students. *American Journal of Preventive Medicine, 12*, 22-30.
- Kolbe, L. J., Kann, L., & Collins, J. L. (1993). Overview of the Youth Risk Surveillance System. *Public Health Reports, 108*, 2-10.
- Korn, M. L., Botsis, A. J., Kotler, M., Plutchik, R., Conte, H. R., Finkelstein, G., Grosz, D., Kay, S., Brown, S. L., & van Praag, H. M. (1992). The suicide and aggression survey: A semistructured instrument for the measurement of suicidality and aggression. *Comprehensive Psychiatry, 33*, 359-365.
- Kupersmidt, J. B., & Coie, J. D. (1990). Preadolescent peer status, aggression, and school adjustment as predictors of externalizing problems in adolescence. *Child Development, 61*, 1350-1362.
- Lardon, C., & Jason, L. A. (1992). Validating a brief pupil evaluation inventory. *Journal of Abnormal Child Psychology, 20*, 367-376.
- Maguire, K., & Pastore, A. L. (1998). *Sourcebook of criminal justice statistics—1997* (NCJ-171147). Washington, DC: U.S. Department of Justice.
- Matthews, K. A., & Angulo, J. (1980). Measurement of the Type A behavior pattern in children: Assessment of children's competitiveness, impatience-anger, and aggression. *Child Development, 51*, 466-475.
- Moffitt, T. E., & Silva, P. A. (1988). Self-reported delinquency, neuropsychological deficit, and history of attention deficit disorder. *Journal of Abnormal Child Psychology, 16*(5), 553-569.
- Morrison, E. F. (1993). The measurement of aggression and violence in hospitalized psychiatric patients. *International Journal of Nursing Studies, 30*, 51-64.
- Olweus, D. (1979). Stability of aggressive reaction patterns in males: A review. *Psychological Bulletin, 86*, 852-875.
- Orpinas, P., Basen-Engquist, K., Grunbaum, J. A., & Parcel, G. S. (1995). The co-morbidity of violence-related behaviors with health-risk behaviors in a population of high school students. *Journal of Adolescent Health, 16*, 216-225.
- Orpinas, P., Kelder, S., Frankowski, R., Murray, N., Zhang, Q., & McAlister, A. (2000). Outcome evaluation of a multi-component violence-prevention program for middle school students: The Students for Peace project. *Health Education Research, 15*(1), 45-58.
- Orpinas, P., Murray, N., & Kelder, S. (1999). Parental influences on students' aggressive behavior and weapon-carrying. *Health Education and Behavior, 26*(6), 774-787.
- Orpinas, P., Parcel, G. S., McAlister, A., & Frankowski, R. (1995). Violence-prevention in middle schools. A pilot evaluation. *Journal of Adolescent Health, 17*, 360-374.
- Oyserman, D., & Saltz, E. (1993). Competence, delinquency, and attempts to attain possible selves. *Journal of Personality and Social Psychology, 65*, 360-374.
- Patterson, G. R., & Stouthamer-Loeber, M. (1984). The correlation of family management practices and delinquency. *Child Development, 55*, 1299-1307.
- Plutchik, R., & van Praag, H. M. (1990). A self-report measure of violence risk, II. *Comprehensive Psychiatry, 31*, 450-456.
- Tremblay, R. E., Pihl, R. O., Vitaro, F., & Dobkin, P. L. (1994). Predicting early onset of male antisocial behavior from preschool behavior. *Archives of General Psychiatry, 51*, 732-739.
- Valois, R. F., Vincent, M. L., McKeown, R. E., Garrison, C. Z., & Kirby, S. D. (1993). Adolescent risk behaviors and the potential for violence: A look at what's coming to campus. *College Health, 41*, 141-147.
- Weis, J. G. (1986). Issues in the measurement of criminal careers. In A. Blumstein, J. Cohen, J. A. Roth, & C. A. Visher (Eds.), *Criminal careers and "career criminals"* (Vol. II, pp. 1-51). Washington, DC: National Academy Press.

Yudofsky, S. C., Silver, J. M., Jackson, W., Endicott, J., & Williams, D. (1986). The overt Aggression Scale for the objective rating of verbal and physical aggression. *American Journal of Psychiatry*, *143*, 35-39.

Requests for reprints should be addressed to Pamela Orpinas, Department of Health Promotion and Behavior, University of Georgia, 300 River Road, Athens, GA 30602; e-mail: porpinas@coe.uga.edu.