Validity of DSM-IV Conduct Disorder in 4½–5-Year-Old Children: A Longitudinal Epidemiological Study

Julia Kim-Cohen, Ph.D.

Louise Arseneault, Ph.D.

Avshalom Caspi, Ph.D.

Mónica Polo Tomás, B.Sc.

Alan Taylor, M.A., M.Sc.

Terrie E. Moffitt, Ph.D.

Objective: This longitudinal study of a nonreferred, population-based sample tested the concurrent, convergent, and predictive validity of DSM-IV conduct disorder in children 4½–5 years of age.

Method: In the Environmental Risk Longitudinal Twin Study, a representative birth cohort of 2,232 children, the children's mothers were interviewed and the teachers completed mailed questionnaires to assess the children's past 6-month conduct disorder symptoms. Children with three or more symptoms were diagnosed with conduct disorder, and a subset with five or more symptoms was diagnosed with "moderate-to-severe" conduct disorder.

Results: The prevalence of conduct disorder and moderate-to-severe conduct disorder were 6.6% and 2.5%, respectively. Children diagnosed with conduct disorder were significantly more likely than comparison subjects to self-report antisocial behaviors, to behave disruptively during observational assessment, and to have risk factors known to be associated with conduct disorder in older children (effect sizes ranging from 0.26 to 1.24). Five-year-olds diagnosed with conduct disorder were significantly more likely than comparison subjects to have behavioral and educational difficulties at age 7. Increased risk for educational difficulties at age 7 persisted after control for IQ and attention deficit hyperactivity disorder diagnosis at age 5.

Conclusions: Behavioral problems of preschool-age children meeting diagnostic criteria for conduct disorder should not be ignored. Appropriate intervention should be provided to prevent ongoing behavioral and academic problems.

Childhood conduct disorder is a top mental health priority (1). Evidence from prospective longitudinal data shows that childhood conduct disorder precedes a variety of major axis I psychiatric disorders (2), suggesting that treating conduct disorder might significantly reduce the burden of adult mental disorder. Preschool intervention is desirable to prevent chronic conduct disorder (3). To intervene early, valid methods must be available to diagnose conduct disorder in young children, and in this *Journal*, Keenan and Wakschlag (4) called for studies of the validity of conduct disorder diagnoses in children under age 6. This article responds, reporting the validity of applying the DSM-IV conduct disorder diagnosis to a representative, nonreferred sample of over 2,000 4½–5-year-old boys and girls.

The validity of the DSM-based conduct disorder diagnosis in very young children can be ascertained by testing its concurrent, convergent, and predictive validity (5). In the present study, the diagnosis of conduct disorder would have concurrent validity if children with and without a conduct disorder diagnosis differed on other, independent measures of conduct problems. The diagnosis would have convergent validity if children with and without a conduct disorder diagnosis differed on measures of risk factors that are known to correlate with conduct disorder in older children. Risk factors associated with conduct

disorder in older children include male sex, sociodemographic characteristics (e.g., low income, teen motherhood), quality of family environment (e.g., domestic violence, low social support), parenting (e.g., child harm), parental psychopathology, and co-occurring behavioral and neurocognitive child characteristics (e.g., hyperactivity, low IQ) (6). The diagnosis would have predictive validity if children with and without a conduct disorder diagnosis at age 5 years differed on measures of behavioral and educational functioning later in childhood.

Additionally, we explored whether any children who were diagnosed with conduct disorder at age 5 did not have any conduct disorder symptoms 2 years later and whether these "remitted" children had outcomes at age 7 that would warrant concern.

Method

Participants

The participants were members of the Environmental Risk Longitudinal Twin (E-Risk) Study. The E-Risk Study sampling frame was two consecutive birth cohorts (1994 and 1995) in a birth register of twins born in England and Wales (7). Of the 15,906 twin pairs born in these 2 years, 71% joined the register.

The E-Risk Study probability sample was drawn with a high-risk stratification sampling procedure. High-risk families were those in which the mother had her first birth when she was 20 years of age or younger. We used this sampling 1) to replace high-risk families

who were selectively lost to the register through nonresponse and 2) to ensure sufficient base rates of problem behavior, given the low base rates expected for young children. Age at first childbearing was used as the risk-stratification variable because it was present for virtually all families in the register, it is relatively free of measurement error, and early childbearing is a known risk factor for children's problem behaviors (8, 9). The sampling strategy resulted in a final sample in which one-third of the study mothers (younger only, N=314) constituted a 160% oversample of mothers who were at high risk based on their young age at first birth (15–20 years). The other two-thirds of study mothers (N=802) accurately represented all mothers in the general population (ages 15–48) in England and Wales in 1994–1995 (estimates derived from the General Household Survey [10]). To provide unbiased statistical estimates that can be generalized to the population of British families with children born in the 1990s, the data reported in this article were corrected with weighting to represent the proportion of young mothers in that population.

The E-Risk Study sought a sample size of 1,100 families to allow for attrition in future years of the longitudinal study while retaining statistical power. An initial list of families who had same-sex twins was drawn from the register to target for home visits, with a 10% oversample to allow for nonparticipation. Of the families from the initial list, 1,116 (93%) participated in home-visit assessments when the twins were age 5 years, forming the base sample for the study: 4% of the families refused, and 3% could not be reached after many attempts. Written informed consent was obtained from the mothers. With the parents' permission, questionnaires were mailed to the children's teachers, and the teachers returned questionnaires for 94% of the cohort children.

A follow-up home visit was conducted 18–24 months after the children's assessment at age 5 (hereafter called the age-7 follow-up). Follow-up data were collected for 98% of the 1,116 E-Risk Study families. At this follow-up, teacher questionnaires were obtained for 91% of the 2,232 E-Risk Study children (93% of those taking part in the follow-up). The E-Risk Study has received ethical approval from the ethics committee of Maudsley Hospital.

Conduct Disorder Diagnosis at Age 5

We derived a research diagnosis of the children's conduct disorder based on the mothers' and the teachers' reports on 14 of 15 DSM-IV symptoms of conduct disorder. The symptoms covered fighting, bullying, lying, stealing, cruelty to people or animals, vandalism, and rule violations. The "forced sexual activity" symptom was excluded as inappropriate for 5-year-olds. The mothers' reports were obtained in a face-to-face standardized interview in the family home, and the teachers' reports were obtained through mail. A child was considered to have a given symptom if either the mother or the teacher reported the symptom as being "very true or often true" of the child over the past 6 months at 4½–5 years of age. We counted a symptom as present if reported by either source from evidence that this approach enhances diagnostic validity (11, 12). Symptom counts ranged from 0 to 11. Consistent with DSM-IV criteria, children with 3 or more symptoms were assigned a conduct disorder diagnosis. A subset of children with 5 or more symptoms was assigned a "moderate-to-severe" conduct disorder diagnosis.

Validity Measures

Distributions and prevalence rates in this cohort for measures used in concurrent, convergent, and predictive validity analyses (Table 1) have been reported elsewhere (9). (An appendix with detailed descriptions is available on request from the first author.)

Statistical Analysis

To provide unbiased statistical estimates that could be generalized to the population of British families with children born in the 1990s, the data reported in this article were corrected with weighting to represent the proportion of young mothers in that population (35).

Group differences were evaluated with odd ratios (for dichotomous variables) and t tests (for continuous variables). We calculated the effect sizes of the obtained group differences from the following formula:

$$d=(\text{Mean}_{\text{conduct disorder}} - \text{Mean}_{\text{no conduct disorder}})/\text{sample standard deviation}$$

For dichotomous variables, we estimated the standardized mean difference statistic (d) by taking the product of the log odds ratio and (square root)$3/\pi$ (36). Operationally defined, d=0.2 is a small effect size, d=0.5 is a medium effect size, and d=0.8 is a large effect size (37).

Statistical analysis was complicated by the fact that our twin study contained two children from each family. Thus, we analyzed data with tests based on the sandwich or Huber/White variance estimator (38, 39) with Stata 7.0 (40), which adjusts estimated standard errors to account for the dependence in the data because of analyzing two children from the same family.

Results

Prevalence and Sex Differences

The prevalence of conduct disorder, weighted to represent the population (reported Ns are unweighted), was 6.6% (N=189) in the total sample, 9.9% (N=136) for the boys, and 3.5% (N=53) for the girls. The prevalence of "moderate-to-severe" conduct disorder, weighted to represent the population, was 2.5% (N=75) in the total sample, 4.2% (N=62) for the boys, and 0.9% (N=13) for the girls. The boys' risk of a conduct disorder diagnosis at age 5 years was 3–5 times that of the girls (odds ratio=3.0, 95% confidence interval [CI]=2.0–4.5 for conduct disorder; odds ratio=4.9, 95% CI=2.4–9.9 for moderate-to-severe conduct disorder).

Three findings are notable about children diagnosed with conduct disorder at age 5 (Table 2). First, symptoms of aggression, theft, and destructive behavior were common, whereas rule violations were rare. Second, there were few significant sex differences in the prevalence of conduct disorder symptoms. Third, where tendencies toward sex differences were observed, conduct-disordered girls' patterns of symptoms reflected relational aggression (e.g., bullying or threatening people, cruelty, lying, or cheating), whereas conduct-disordered boys' pattern of symptoms reflected conventional conduct disorder symptoms (e.g., hitting others, starting fights, destroying property, stealing, and breaking into disallowed places).

Concurrent Validity

Table 3 shows the conduct disorder group was significantly more likely than the comparison group to self-report antisocial behaviors in the Berkeley Puppet Interview (13, 14), including hitting, lying, swearing, fighting, stealing, and breaking things. Observational measurements revealed that the conduct disorder group was also significantly more disruptive than the comparison group during a competitive game of Snap! (15).

TABLE 1. Description of Measures Included in Concurrent, Convergent, and Predictive Validity Analyses of DSM-IV Conduct Disorder in Children at Age 5

Measure	Number of Items	Range	Proportion (%)	Interrater/ Test-Retest Reliability (r)	Internal Consistency (alpha)
Children's conduct problems (other than DSM-IV) at age 5					
Self-report: Berkeley Puppet Interview	19	31 to 106		0.90	
Observational: Snap! Disruptive Behavior Scale (standardized)	4	−1 to 6		0.83	
Sociodemographic risks					
Mother has no educational qualification	1		13		
Manual occupational status	1		34		
Yearly household income ≤£10,000	1		12		
Teen motherhood	1		28		
Family environment risks					
Single parent for past 5 years	1		3		
2 or more address changes in past 5 years	1		17		
Physical partner abuse in past 5 years	24	0 to 40		0.77	0.89
Mother's perceived social support	12	0 to 24			0.76
Parenting risks					
Child harm			12	0.90	
Maternal expressed emotion: negativity		0 to 5		0.84	
Parental psychopathology					
Maternal depression in past 5 years			26	0.90	
Maternal antisocial behavior	39	0 to 60			0.90
Biological father's antisocial behavior	44	0 to 88		0.75	0.95
Child characteristics at age 5					
Neurocognitive predictors					
IQ (Wechsler Preschool and Primary Scale of Intelligence— Revised, vocabulary and block design)	two subtests	52 to 145		0.89	
Executive function, Day-Night Task, Sentence Working Memory Task, Wechsler Preschool and Primary Scale of Intelligence—Revised, mazes	three subtests	1.5 to 20		0.84	
Co-occurring behavioral predictors					
Mother's report of children's behavioral syndromes					
Impulsivity scale	4	0 to 8			0.69
Hyperactivity scale	4	0 to 8			0.82
Inattention scale	9	0 to 18			0.85
Attention deficit hyperactivity disorder (ADHD) symptom scale	18	0 to 34			0.90
Internalizing scale	31	0 to 44			0.84
Prosocial behavior scale	10	0 to 20			0.76
Teacher's report of children's behavioral syndromes					
Impulsivity scale	4	0 to 8			0.85
Hyperactivity scale	5	0 to 8			0.86
Inattention scale	9	0 to 18			0.92
ADHD symptom scale	18	0 to 34			0.94
Internalizing scale	35	0 to 50			0.85
Prosocial behavior scale	10	0 to 20			0.92
Observational measure: Snap! Hyperactive/Inattentive Scale (standardized)	3	−2 to 4		0.83	
Children's behavioral outcomes at age 7					
Mother's report of children's behavioral syndromes					
ADHD symptom scale	17	0 to 34			0.91
Aggression scale	19	0 to 38			0.89
Delinquency scale	11	0 to 18			0.60
Internalizing scale	31	0 to 43			0.86
Prosocial behavior scale	10	0 to 20			0.80
Teacher's report of children's behavioral syndromes					
ADHD symptom scale	17	0 to 34			0.94
Aggression scale	25	0 to 48			0.94
Delinquency scale	9	0 to 48			0.72
Internalizing scale	35	0 to 46			0.87
Prosocial behavior scale	10	0 to 20			0.93
Educational outcomes at age 7					
Referred for special education	1		22		
Standard reading score (Test of Word Reading Efficiency) below 90			9	0.96	
English school performance below average	1		31		
Math school performance below average	1		27		

Reference Citations
Ablow et al., 1999 (13); Measelle et al., 1998 (14)
Hughes et al., 2002 (15)
Jowell et al., 1997 (16)
Jowell et al., 1997 (16)
Jowell et al., 1997 (16)
Caspi et al., 1996 (17)
Caspi et al., 1996 (17)
Moffitt et al., 1997 (18); Straus, 1990 (19)
Simons and Johnson, 1996 (20)
Dodge et al., 1995 (21)
Caspi et al., 2004 (22)
Robins et al., 1995 (23)
Achenbach, 1997 (24)
Achenbach, 1997 (24); Caspi et al., 2002 (25)
Wechsler, 1990 (26); Sattler, 1992 (27)
Daneman and Carpenter, 1980 (28); Gerstadt et al., 1994 (29); Wechsler, 1990 (26)
Achenbach, 1991 (30); Goodman, 1994 (31); Sclare, 1997 (32)
Achenbach, 1991 (33)
Hughes et al., 2002 (15)
Achenbach, 1991 (30); Goodman, 1994 (31); Sclare, 1997 (32)
Achenbach, 1991 (33)
Achenbach, 1991 (33)
Torgesen et al., 1999 (34)
Achenbach, 1991 (33)
Achenbach, 1991 (33)

Convergent Validity

Table 4 presents risk factors known to correlate with conduct disorder in older children. Risk factors in every domain (e.g., sociodemographic, family environment, parenting, parental psychopathology, and child characteristics) differed significantly between the conduct disorder and the comparison groups. All measures of children's co-occurring behavioral problems, based on the mothers' and the teachers' reports, discriminated between children with and without an age-5 conduct disorder diagnosis, except teacher-reported internalizing behavior. The largest effect sizes were observed for the children's behaviors related to attention deficit hyperactivity disorder (ADHD). Our observational measure agreed that the conduct disorder group was rated significantly more hyperactive or inattentive than the comparison group during the Snap! game. The risk factors related to parenting and parental psychopathology were also among the strongest group discriminators (with effect sizes ranging from 0.51 to 0.85).

Predictive Validity

The children meeting the criteria for a conduct disorder diagnosis at age 5 years were significantly more likely than the comparison subjects to have higher levels of age-7 mother- and teacher-reported ADHD symptoms, aggression, and delinquency and lower levels of prosocial behavior (Table 5). Relative to comparison subjects, the conduct disorder group was 2–4 times more likely to have age-7 educational problems. After we controlled for age-5 IQ and ADHD diagnosis, the conduct disorder group continued to be at an increased risk for poorer reading (adjusted odds ratio=2.0, 95% CI=1.2–3.2), English (adjusted odds ratio=1.8, 95% CI=1.2–2.8), and math (adjusted odds ratio=1.6, 95% CI=1.0–2.4) skills. Sixty-five percent of the children with a conduct disorder diagnosis and 80% of the children with a moderate-to-severe conduct disorder diagnosis at age 5 years had one or more educational problems at age 7 years.

Continuity or Change?

Compared to undiagnosed children, the age-5 conduct disorder group was at a significantly greater risk for a conduct disorder diagnosis at age 7 years (odds ratio=20.6, 95% CI=12.5–34.1). More boys than girls maintained a conduct disorder diagnosis at both ages (odds ratio=3.8, 95% CI=1.1–12.6).

Despite this continuity, among the 189 5-year-olds who met the diagnostic criteria for conduct disorder, 83 (49% weighted) had no conduct disorder symptoms at age 7 (Figure 1). However, of the subset of 75 children who met the diagnostic criteria for moderate-to-severe conduct disorder at 5 years of age, only 21 (28% weighted) had no conduct disorder symptoms at age 7 years. We compared the group of children who had an age-5 conduct disorder diagnosis and no age-7 conduct disorder symptoms ("remitted" group) against the children who did not have an age-5 con-

TABLE 2. Prevalence of Conduct Disorder Symptoms at Age 5 in Total Sample of Children and in Children Meeting Criteria for a Conduct Disorder Diagnosis[a]

| DSM-IV Conduct Disorder Symptom | Total Sample (N=2,232) | | Children Meeting Conduct Disorder Criteria | | | | | | | |
| | | | Total (N=189) | | Boys (N=136) | | Girls (N=53) | | Analysis of Boys:Girls[b] | |
	N[c]	%	N	%	N	%	N	%	Odds Ratio	95% CI
Aggression toward people and animals										
Bullies or threatens people	91	3.2	82	44.6	53	39.8	29	57.5	0.5	0.2–0.9*
Hits people with dangerous objects	152	5.7	101	52.7	85	61.1	16	30.0	3.8	1.6–9.2**
Is physically cruel toward others	101	3.3	74	36.5	51	33.3	23	45.0	0.6	0.3–1.3
Is physically cruel toward animals	30	0.9	25	11.5	18	12.0	7	10.0	1.3	0.5–3.5
Uses force to take things from others	179	6.7	113	60.1	83	62.0	30	55.0	1.3	0.6–2.9
Starts fights	211	7.5	123	65.3	95	67.6	28	59.0	1.5	0.7–3.2
Destruction of property										
Deliberately sets fires	14	0.5	11	5.4	8	5.6	3	5.0	1.3	0.2–7.3
Deliberately destroys others' property	264	9.6	140	70.3	106	74.1	34	60.0	1.9	0.9–3.9
Deceitfulness or theft										
Breaks into disallowed places	25	1.1	21	14.2	18	16.7	3	7.5	2.8	0.6–12.0
Lies or cheats	160	5.7	90	45.3	59	40.7	31	57.5	0.5	0.2–1.1
Steals	60	2.0	43	21.6	37	25.9	6	10.0	3.1	1.1–8.4*
Serious violations of rules										
Stays out at night too late	4	0.1	2	0.7	2	0.9	0	0.0	—	
Runs away from home	5	0.3	4	3.4	3	2.8	1	5.0	0.6	0.1–6.8
Skips school or is a truant	12	0.3	7	2.7	4	1.9	3	5.0	0.5	0.1–3.6

[a] Tables 2–6 present means and percents for children with a conduct disorder diagnosis, but in all cases, scores were more extreme for the subset of children with "moderate-to-severe" conduct disorder. Details are available from the first author.
[b] Based on estimates of standard errors adjusted to account for the dependence in the data because of analyzing two children in the same family (38, 39).
[c] Ns are unweighted; proportions are weighted to represent the population of British families.
*p<0.05.   **p<0.01.

TABLE 3. Concurrent Validity: Comparison of Children With and Without a Conduct Disorder Diagnosis at Age 5 on Measures of Conduct Problems Other Than DSM-IV Diagnosis

| Item Number | Measure | Total Score | | | | Group Difference[a] | | Effect Size (d)[b] |
| | | Children With Conduct Disorder (N=147) | | Children Without Conduct Disorder (N=1,733) | | | | |
		Mean	SD	Mean	SD	t	df	
	Berkeley Puppet Interview of Children's Self-Reported Conduct Problems (item range=1–7)							
		59.0	15.6	51.0	13.0	5.54***	971	0.56
1	I lose my temper.	3.5	1.9	3.6	2.0	0.71	974	0.05
2	I take things that don't belong to me.	3.4	1.8	3.0	1.7	2.08*	970	0.23
3	It's funny when a kid gets in trouble at school.	3.7	2.0	2.9	1.7	3.94***	971	0.43
4	I steal.	2.9	1.7	2.4	1.3	2.46*	968	0.33
5	I tease other kids.	3.5	1.9	2.8	1.6	3.76***	971	0.40
6	I tell lies.	3.2	1.8	2.7	1.5	2.61**	972	0.30
7	I hit kids a lot.	3.0	1.7	2.5	1.3	2.97**	969	0.33
8	It's fun to tease other kids.	2.9	1.7	2.5	1.3	2.74**	972	0.26
9	I like to mess up other kids' games or work.	3.0	1.7	2.6	1.4	2.19*	967	0.26
10	I'm nasty to animals.	2.9	1.6	2.4	1.2	3.13**	971	0.35
11	I fight with other kids a lot.	3.1	1.8	2.6	1.4	3.49***	971	0.31
12	I hit my mommy or daddy.	3.1	1.7	2.6	1.3	2.84**	970	0.33
13	I yell at mommy or daddy.	3.0	1.7	2.6	1.4	2.27*	971	0.26
14	I break other people's things.	2.7	1.5	2.4	1.2	2.38*	969	0.22
15	I start fires.	2.8	1.6	2.4	1.2	3.12**	969	0.28
16	I don't do what mommy/daddy asks me to do.	3.1	1.8	2.9	1.6	1.38	966	0.12
17	I don't do what my teacher asks me to do.	3.0	1.7	2.5	1.3	2.74**	967	0.33
18	I swear or say bad words.	3.0	1.7	2.5	1.3	3.07**	969	0.33
19	I cheat when playing a game.	3.1	1.7	2.8	1.6	1.57	970	0.18
	Snap! Observational Measure: Disruptive Behavior Scale (standardized)[c]	0.30	1.2	−0.05	0.96	3.66***	1058	0.32

[a] Standard errors and test statistics include adjustments for the dependence in the data because of analyzing two children in the same family (38, 39). Thus, the degrees of freedom are based on the number of families rather than the number of children. Analyses are based on all available cases, and therefore, different Ns across Tables 3, 4, and 5 reflect different amounts of missing data.
[b] Differences between groups can be interpreted in terms of standard deviation units (d), where d=0.2 is considered a small effect size, d=0.5 is a medium effect size, and d=0.8 is a large effect size (37).
[c] For this measure, N=180 for children with conduct disorder and N=1,931 for children without conduct disorder.
*p<0.05.   **p<0.01.   ***p<0.001.

TABLE 4. Convergent Validity: Comparison of Children With and Without a Conduct Disorder Diagnosis at Age 5 on Risk Factors Known to Be Associated With Conduct Disorder in Older Children

Associated Risks at Age 5	Children With Conduct Disorder (N=189)			Children Without Conduct Disorder (N=2,043)			Group Difference[a]		Effect Size (d)[b]
	N[c]	Mean or %	SD	N	Mean or %	SD	Odds Ratio or t	95% CI or df	
Sociodemographic risks									
Mother has no educational qualification	63	25.8%		343	12.3%		2.5***	1.7–3.7	0.51
Manual occupational status	129	57.5%		819	31.9%		2.9***	1.9–4.3	0.59
Yearly household income ≤£10,000	64	28.4%		322	11.2%		3.1***	2.1–4.8	0.62
Teen motherhood	140	53.0%		984	26.4%		3.1***	2.1–4.7	0.62
Family environment risks									
Single parent for past 5 years	16	6.6%		76	2.9%		2.4*	1.2–4.9	0.48
Two or more address changes in past 5 years	57	26.5%		371	16.6%		1.8**	1.2–2.7	0.32
Physical partner abuse in past 5 years		5.5	7.7		2.6	5.5	4.54***	1094	0.43
Mother's perceived social support		13.2	6.1		15.6	5.4	3.96***	1114	0.42
Parenting risks									
Maternal expressed emotion: negativity		2.3	1.2		1.4	0.9	9.00***	1010	0.85
Child harm is probable or definite	63	28.9%		244	11.2%		3.2***	2.1–5.0	0.64
Parental psychopathology									
Maternal depression in past 5 years	84	45.0%		550	24.8%		2.5***	1.7–3.7	0.51
Maternal antisocial behavior		20.9	14.6		10.6	8.9	7.39***	1112	0.85
Biological father's antisocial behavior		27.4	20.2		13.9	15.6	7.19***	1108	0.75
Child characteristics									
Neurocognitive predictors									
IQ		91.5	13.8		98.3	14.3	5.47***	1112	0.48
Executive function		11.0	3.2		11.8	3.0	2.66**	1112	0.26
Co-occurring behavioral syndromes									
Mother's report									
Impulsivity scale		5.6	2.0		3.3	2.1	13.88***	1114	1.12
Hyperactivity scale		4.9	2.9		2.1	2.3	11.53***	1114	1.07
Inattention scale		8.7	5.1		4.4	4.0	10.17***	1114	0.94
Attention deficit hyperactivity disorder (ADHD) symptom scale		19.2	8.3		9.7	7.0	13.64***	1114	1.24
Internalizing scale		13.3	8.8		8.0	6.4	6.45***	1114	0.69
Prosocial behavior scale		14.9	3.6		16.4	3.2	4.96***	1114	0.44
Teacher's report									
Impulsivity scale		2.8	2.5	1.1	1.7		7.30***	1051	0.80
Hyperactivity scale		2.4	2.5	0.9	1.6		6.13***	1051	0.72
Inattention scale		5.5	5.2	2.6	3.7		5.58***	1053	0.64
ADHD symptom scale		10.7	9.2	4.6	6.1		6.58***	1052	0.78
Internalizing scale		6.9	6.5	5.8	5.7		1.88	1050	0.18
Prosocial behavior scale		9.1	5.2	11.9	4.8		5.40***	1028	0.56
Attention deficit hyperactivity disorder (ADHD) diagnosis	46	24.3%		92	4.4%		7.0***	4.4–11.3	1.07
Snap! Hyperactive/Inattentive Scale (standardized)		0.27	1.0		−0.02	1.0	3.11**	1058	0.29

[a] Continuous variables were analyzed with t tests and their degrees of freedom and categorical variables with odds ratios and their 95% confidence intervals (CIs). Standard errors, 95% CIs, and test statistics include adjustments for the dependence in the data because of analyzing two children in the same family (38, 39). Thus, degrees of freedom are based on the number of families rather than the number of children.
[b] Differences between groups can be interpreted in terms of standard deviation units (d), where d=0.2 is considered a small effect size, d=0.5 is a medium effect size, and d=0.8 is a large effect size (37).
[c] Ns are unweighted; proportions are weighted to represent the population of British families.
*p<0.05.   **p<0.01.   ***p<0.001.

duct disorder diagnosis (comparison group) on age-7 outcome measures (Table 6). The findings suggested that the "remitted" children continued to experience clinically significant behavioral and academic difficulties 2 years later, despite their apparent remission from conduct disorder.

Discussion

The utility of a diagnostic framework lies primarily in its ability to identify individuals who may need treatment. Applying DSM-IV diagnostic criteria for conduct disorder to 4½–5-year-old children appears to have succeeded in

doing that. In a nonreferred, population-based sample, the conduct disorder diagnosis classified a small subset of children who were aggressive and antisocial. They also had lower executive functioning, lower IQ, and co-occurring ADHD symptoms—neurodevelopmental hallmarks of life-course persistent antisocial behavior (41). Moreover, diagnosed children were disproportionately likely to come from backgrounds marked by low social class, single parenthood, family disruption, and parental psychopathology. Conduct-disordered children in our sample were disproportionately likely to have experienced harsh

TABLE 5. Predictive Validity: Comparison of Children With and Without a Conduct Disorder Diagnosis at Age 5 on Outcome Measures at Age 7

Child Functioning at Age 7	Children With Conduct Disorder (N=184)		Children Without Conduct Disorder (N=1,994)		Group Difference[a]		Effect Size[b]
	Mean	SD	Mean	SD	t	df	d
Behavioral outcomes							
Mother's report							
Attention deficit hyperactivity disorder (ADHD) symptom scale	16.4	9.2	8.8	6.8	9.40*	1088	0.94
Aggression scale	17.9	8.5	7.8	5.9	12.63*	1088	1.38
Delinquency scale	4.3	3.1	1.6	1.7	10.21*	1088	1.08
Internalizing scale	11.5	8.1	7.0	5.9	5.90*	1088	0.64
Prosocial behavior scale	14.7	4.1	16.5	3.2	5.24*	1088	0.49
Teacher's report							
ADHD symptom scale	7.8	7.7	4.2	6.1	5.05*	1026	0.52
Aggression scale	8.9	9.0	3.6	5.9	6.75*	1026	0.70
Delinquency scale	1.5	2.1	0.6	1.3	5.18*	1026	0.52
Internalizing scale	6.7	6.3	5.7	6.0	1.86	1026	0.16
Prosocial behavior scale	10.3	4.8	12.9	4.8	5.36*	1023	0.54
	N[c]	%	N	%	Odds Ratio	95% CI	d
Educational outcomes							
Referred for special education	71	38.0	412	20.8	2.3*	1.5–3.5	0.46
Standard reading score below 90	51	23.8	208	7.7	3.7*	2.4–5.8	0.72
English school performance below average	107	55.1	606	28.7	3.0*	2.1–4.4	0.61
Math school performance below average	92	48.2	540	25.5	2.7*	1.9–4.0	0.55

[a] Continuous variables were analyzed with t tests and their degrees of freedom and categorical variables with odds ratios and their 95% confidence intervals (CIs). Standard errors, 95% CIs, and test statistics include adjustments for the dependence in the data because of analyzing two children in the same family (38, 39). Thus, degrees of freedom are based on the number of families rather than the number of children.
[b] Differences between groups can be interpreted in terms of standard deviation units (d), where d=0.2 is considered a small effect size, d=0.5 is a medium effect size, and d=0.8 is a large effect size (37).
[c] Ns are unweighted; proportions are weighted to represent the population of British families.
*p<0.001.

FIGURE 1. The Distribution of Conduct Disorder Symptoms Among Children Meeting Diagnostic Criteria for Conduct Disorder at Age 5
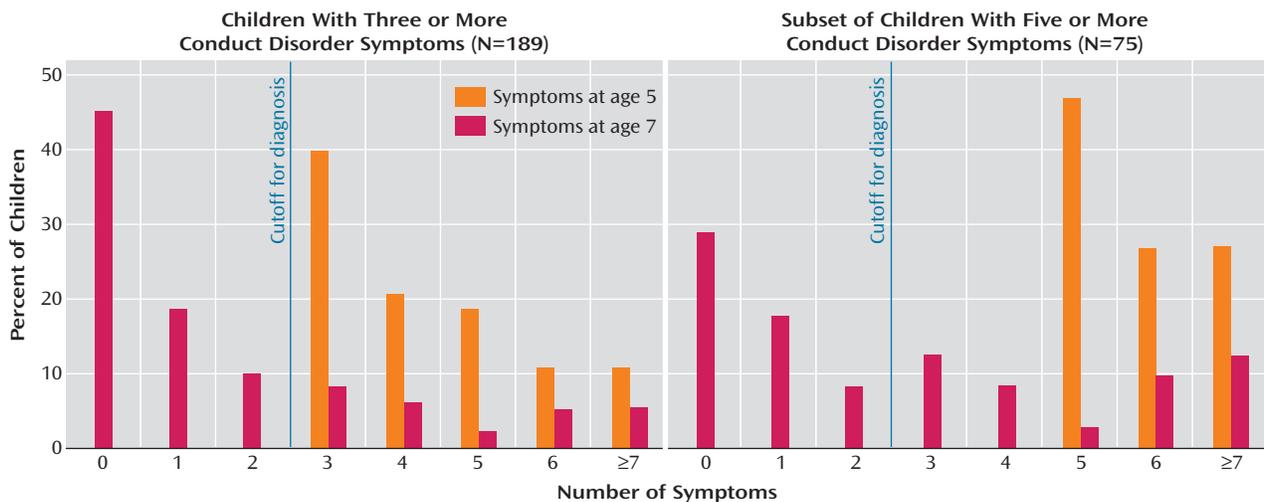
parenting and physical harm. Thus, the life contexts of young children identified as having conduct disorder in our study resemble the "criminogenic" environments that have been shown to interact with neurodevelopmental vulnerabilities to exacerbate behavior problems and promote antisocial personality development (41). Concerns about the welfare of children diagnosed with conduct disorder increase, given that these children, regardless of

whether they had any conduct disorder symptoms at follow-up, were disproportionately likely to have continuing behavioral and educational difficulties. Effect sizes for the association between risk factors and conduct disorder in our sample of young children are comparable to those reported for older children (42, 43).

Whether valid diagnoses can be made of disruptive behavior disorders in preschool-age children has been a fo-

TABLE 6. Predictive Validity: Comparison of Children With a Conduct Disorder Diagnosis at Age 5 and No Conduct Disorder Symptoms at Age 7 ("Remitted") Versus Children With No Conduct Disorder Diagnosis At Age 5 (Comparison Group) on Outcome Measures at Age 7

Child Functioning at Age 7	Diagnostic Grouping						
	"Remitted" Group (conduct disorder at age 5 but no age-7 conduct disorder symptoms) (N=83)		Comparison Group (no age-5 conduct disorder) (N=1,994)		Group Difference[a]		Effect Size[b]
	Mean	SD	Mean	SD	t	df	d
Behavioral outcomes							
Mother's report							
Attention deficit hyperactivity disorder							
(ADHD) symptom scale	12.3	7.9	8.8	6.8	3.51***	1065	0.48
Aggression scale	12.1	5.8	7.8	5.9	5.70***	1065	0.74
Delinquency scale	2.6	1.6	1.6	1.7	5.38***	1065	0.61
Internalizing scale	9.3	7.0	7.0	5.9	2.49*	1065	0.36
Prosocial behavior scale	16.1	3.5	16.5	3.2	1.13	1065	0.12
Teacher's report							
ADHD symptom scale	5.2	5.4	4.2	6.1	1.56	1001	0.17
Aggression scale	6.6	6.9	3.6	5.9	3.43***	1001	0.47
Delinquency scale	0.9	1.4	0.6	1.3	1.91	1001	0.22
Internalizing scale	5.6	5.7	5.7	6.0	0.13	1001	0.02
Prosocial behavior scale	11.2	5.1	12.9	4.8	2.56*	998	0.34
	N[c]	%	N	%	Odds Ratio	95% CI	d
Educational outcomes							
Referred for special education	27	32.2	412	20.8	1.8*	1.1–3.1	0.32
Standard reading score below 90	21	19.4	208	7.7	2.9***	1.6–5.1	0.59
English school performance below average	42	46.3	606	28.7	2.1**	1.3–3.6	0.41
Math school performance below average	38	42.2	540	25.5	2.1**	1.3–3.6	0.41

[a] Continuous variables were analyzed with t tests and their degrees of freedom and categorical variables with odds ratios and their 95% confidence intervals (CI). Standard errors, 95% CIs, and test statistics include adjustments for dependence in the data because of analyzing two children in the same family (38, 39). Thus, degrees of freedom are based on the number of families rather than the number of children.
[b] Differences between groups can be interpreted in terms of standard deviation units (d), where d=0.2 is considered a small effect size, d=0.5 is a medium effect size, and d=0.8 is a large effect size (37).
[c] Ns are unweighted; proportions are weighted to represent the population of British families.
*p<0.05.   **p<0.01.   ***p<0.001.

cus of controversy (4, 44, 45). Some argue that disruptive behaviors in young children should not be pathologized because aggressive, destructive, and defiant behaviors, which characterize conduct disorder, are thought to be common and developmentally normative in the preschool period (46), and most children will outgrow them (47). Others believe that children falsely identified as having conduct disorder may be stigmatized, and unnecessary referral for treatment may waste health care resources. Some argue that the predictive accuracy of conduct problems for future conduct disorder improves only when children are older (48), and applying diagnostic criteria validated for older children and adolescents to young children (i.e., "down-aging") may promote overdiagnosis (44).

Our research demonstrated that the *co-occurrence* of three or more conduct disorder symptoms, constituting a *syndrome* of problems, is rare and signals poor outcome. In this study, a large majority of children who had at least three (65%) or at least five (80%) DSM-IV conduct disorder symptoms at age 5 years, regardless of age-7 conduct disorder symptom status, had at least one educational difficulty 2 years later. Furthermore, the fact that the children's conduct disorder diagnosis continued to predict teacher-rated educational difficulties at follow-up, after we controlled for age-5 IQ and ADHD diagnosis, suggests that the

conduct disorder diagnosis uniquely captures a subset of young children who may have enduring school-related problems and on whom treatment resources would not have been wasted.

Clinicians wishing to minimize the risk of treating "false positives" can adopt a conservative approach by applying more stringent diagnostic criteria for moderate-to-severe conduct disorder. Moreover, diagnostic precision might be improved if specific age-appropriate guidelines are developed (47). However, the present study shows that minimum DSM-IV criteria for a conduct disorder diagnosis were sensitive enough to identify young children who might benefit from intervention. Long-term studies of early-onset antisocial behavior have shown that true recoveries from conduct disorder are extremely rare because although not all conduct disorder children grow up to have antisocial personality disorder, virtually all children with conduct disorder have mental disorders and poor functioning in adulthood (41). Fortunately, evidence has shown that interventions for preventing chronic conduct disorder can be effective if applied early in life (49).

Four limitations temper the findings of this study. First, our sample comprised mostly white twins living in England and Wales. Therefore, our findings may not generalize to other ethnic or racial groups in other countries or to

singletons. However, our estimate of the prevalence of conduct disorder of 2.5%–6.6% and the sex ratio of 3–5 boys per girl are comparable to other epidemiological studies of singletons in the United States (50–52). Second, we could not address the validity of the oppositional defiant disorder diagnosis because it was not assessed. Third, we report findings based on a "research" diagnosis that may differ from typical practice in clinical settings. However, we interviewed mothers face-to-face to assess symptoms and did not rely on self-administered checklists. Fourth, predictive validity analyses were limited to a 2-year follow-up period. Predictive validity for later childhood and adolescence should be tested.

The strengths of this study include features that directly address the limitations of existing research identified by Keenan and Wakschlag (4): i.e., equal representation of preschool-age girls and boys in a large population-based sample, both teacher and mother reports of DSM-IV conduct disorder symptoms, and a longitudinal design with a high retention rate (98%).

Recently, a standardized interview (53) and observational (54) methods for making diagnoses in preschool-age children have been developed. This study adds to those efforts by demonstrating that interviewing mothers and gathering collateral reports from teachers can validly identify young children with DSM-IV conduct disorder.

References

1. Priorities for Prevention Research at NIMH: A Report by the National Advisory Mental Health Council Workgroup on Mental Disorders Prevention Research: NIH Publication 98-4321. Bethesda, Md, National Institutes of Health, 1998

2. Kim-Cohen J, Caspi A, Moffitt TE, Harrington HL, Milne B, Poulton R: Prior juvenile diagnoses in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. Arch Gen Psychiatry 2003; 60:709–717

3. Bennett KJ, Offord DR: Conduct disorder: can it be prevented? Curr Opin Psychiatry 2001; 14:333–337

4. Keenan K, Wakschlag LS: Can a valid diagnosis of disruptive behavior disorder be made in preschool children? Am J Psychiatry 2002; 159:351–358

5. Anastasi A: Psychological Testing, 7th ed. New York, Macmillan, 1996

6. Lahey BB, Moffitt TE, Caspi A: Causes of Conduct Disorder and Juvenile Delinquency. New York, Guilford, 2003

7. Trouton A, Spinath FM, Plomin R: Twins Early Development Study (TEDS): a multivariate, longitudinal genetic investigation of language, cognition, and behavior problems in childhood. Twin Res 2002; 5:444–448

8. Maynard RA: Kids Having Kids: Economic Costs and Social Consequences of Teen Pregnancy. Washington, DC, Urban Institute Press, 1997

9. Moffitt TE, E-Risk Study Team: Teen-aged mothers in contemporary Britain. J Child Psychol Psychiatry 2002; 43:727–742

10. Bennett N, Jarvis L, Rowlands O, Singleton N, Haselden L: Living in Britain: results from the General Household Survey. London, Her Majesty's Stationery Office, 1996

11. Bird HR, Gould MS, Staghezza B: Aggregating data from multiple informants in child psychiatry epidemiological research. J Am Acad Child Adolesc Psychiatry 1992; 31:78–85

12. Piacentini JC, Cohen P, Cohen J: Combining discrepant diagnostic information from multiple sources: are complex algorithms better than simple ones? J Abnorm Child Psychol 1992; 20:51–63

13. Ablow JC, Measelle JR, Kraemer HC, Harrington R, Luby J, Smider N, Dierker L, Clark V, Dubicka B, Heffelfinger A, Essex MJ, Kupfer DJ: The MacArthur Three-City Outcome Study: evaluating multi-informant measures of young children's symptomatology. J Am Acad Child Adolesc Psychiatry 1999; 38: 1580–1590

14. Measelle JR, Ablow JC, Cowan PA, Cowan CP: Assessing young children's views of their academic, social, and emotional lives: an evaluation of the self-perception scales of the Berkeley Puppet Interview. Child Dev 1998; 69:1556–1576

15. Hughes C, Oksanen H, Taylor A, Jackson J, Murray L, Caspi A, Moffitt TE: "I'm gonna beat you!" SNAP! an observational paradigm for assessing young children's disruptive behavior in competitive play. J Child Psychol Psychiatry 2002; 43:507–516

16. Jowell R, Cutice J, Park A, Brook L, Thomson K, Bryson C: British Social Attitudes: The 14th Report. Aldershot, England, Ashgate, 1997

17. Caspi A, Moffitt TE, Thornton A, Freedman D, Amell JW, Harrington H, Smeijers J, Silva PA: The Life History Calendar: a research and clinical assessment method for collecting retrospective event-history data. Int J Methods Psychiatr Res 1996; 6:101–114

18. Moffitt TE, Caspi A, Krueger RF, Magdol L, Margolin G, Silva PA, Sydney R: Do partners agree about abuse in their relationship? a psychometric evaluation of interpartner agreement. Psychol Assess 1997; 9:47–56

19. Straus MA: Measuring intrafamily conflict and violence: the Conflict Tactics (CT) Scales, in Physical Violence in American Families: Risk Factors and Adaptations to Violence in 8,145 Families. Edited by Straus MA, Gelles RJ. New Brunswick, NJ, Transaction, 1990, pp 403–424

20. Simons RL, Johnson C: The impact of marital and social network support on quality of parenting, in Handbook of Social Support and the Family. Edited by Pierce GR, Sarason BR, Sarason IG. New York, Plenum, 1996, pp 269–287

21. Dodge KA, Pettit GS, Bates JE, Valente E: Social information-processing patterns partially mediate the effect of early physical abuse on later conduct problems. J Abnorm Psychol 1995; 104:632–643

22. Caspi A, Moffitt TE, Morgan J, Rutter M, Taylor A, Arseneault L, Tully L, Jacobs C, Kim-Cohen J, Polo-Tomas M: Maternal expressed emotion predicts children's antisocial behavior problems: using monozygotic-twin differences to identify environmental effects on behavioral development. Dev Psychol 2004; 40:149–161

23. Robins LN, Cottler L, Bucholtz K, Compton W: National Institute of Mental Health Diagnostic Interview Schedule for DSM-IV. St Louis, Washington University School of Medicine, 1995

24. Achenbach TM: Manual for the Young Adult Self-Report and Young Adult Behavior Checklist. Burlington, University of Vermont, Department of Psychiatry, 1997

25. Caspi A, Taylor A, Smart M, Jackson J, Tagami S, Moffitt TE: Can women provide reliable information about their children's fathers? cross-informant agreement about men's lifetime antisocial behavior. J Child Psychol Psychiatry 2002; 42:915–920

26. Wechsler D: Wechsler Preschool and Primary Scale of Intelligence—Revised. London, Psychological Corp (Harcourt, Brace), 1990

27. Sattler JM: Assessment of Children: WISC-III and WPPSI-R Supplement. San Diego, Jerome M Sattler, 1992

28. Daneman M, Carpenter PA: Individual differences in working memory and reading. J Verbal Learning and Verbal Behavior 1980; 19:450–466

29. Gerstadt CL, Hong YJ, Diamond A: The relationship between cognition and action: performance of children 3.5–7 years old on a Stroop-like day-night test. Cognition 1994; 53:129–153

30. Achenbach TM: Manual for the Child Behavior Checklist/4-18 and 1991 Profile. Burlington, University of Vermont, Department of Psychiatry, 1991

31. Goodman R: A modified version of the Rutter Parent Questionnaire including extra items on children's strengths: a research note. J Child Psychol Psychiatry 1994; 35:1483–1494

32. Sclare I: The Child Psychology Portfolio. Windsor, Berkshire, UK, National Foundation for Educational Research-Nelson, 1997

33. Achenbach TM: Manual for the Teacher's Report Form and 1991 Profile. Burlington, University of Vermont, Department of Psychiatry, 1991

34. Torgesen JK, Wagner RK, Rashotte CA: Test of Word Reading Efficiency. Austin, Tex, PRO-ED, 1999

35. Birth Statistics: Series FM1, Number 23. London, Her Majesty's Stationery Office, 1996

36. Haddock CK, Rinkdskopf D, Shadish WR: Using odds ratios as effect sizes for meta-analysis of dichotomous data: a primer on methods and issues. Psychol Methods 1998; 3:339–353

37. Cohen J: A power primer. Psychol Bull 1992; 112:155–159

38. Rogers WH: Regression standard errors in clustered samples. Stata Technical Bull 1993; 13:19–23

39. Williams RL: A note on robust variance estimation for cluster-correlated data. Biometrics 2000; 56:645–646

40. Stata Statistical Software, Release 7.0. College Station, Tex, Stata Corp, 2003

41. Moffitt TE: Life-course-persistent and adolescence-limited antisocial behavior: a 10-year research review and research agenda, in Causes of Conduct Disorder and Juvenile Delinquency. Edited by Lahey BB, Moffitt TE, Caspi A. New York, Guilford, 2003, pp 49–75

42. Nigg JT, Huang-Pollock CL: An early-onset model of the role of executive functions and intelligence in conduct disorder/delin-

quency, in Causes of Conduct Disorder and Juvenile Delinquency. Edited by Lahey BB, Moffitt TE, Caspi A. New York, Guilford, 2003, pp 227–247

43. Rowe R, Maughan B, Pickles A, Costello EJ, Angold A: The relationship between DSM-IV oppositional defiant disorder and conduct disorder: findings from the Great Smoky Mountains Study. J Child Psychol Psychiatry 2002; 43:365–373

44. McClellan JM, Speltz ML: Psychiatric diagnosis in preschool children (letter). J Am Acad Child Adolesc Psychiatry 2003; 42:127–128

45. Wilens TE, Biederman J, Spencer TJ, Monuteaux M: Psychiatric diagnosis in preschool children: Dr Wilens et al reply (letter). J Am Acad Child Adolesc Psychiatry 2003; 42:128–129

46. Tremblay RE, Japel C: Prevention during pregnancy, infancy and the preschool years, in Early Prevention of Adult Antisocial Behaviour. Edited by Farrington DP, Coid JW. Cambridge, UK, Cambridge University Press, 2003, pp 205–242

47. Campbell SB: Behavior Problems in Preschool Children: Clinical and Developmental Issues, 2nd ed. New York, Guilford, 2002

48. Bennett KJ, Offord DR: Screening for conduct problems: does the predictive accuracy of conduct disorder symptoms improve with age? J Am Acad Child Adolesc Psychiatry 2001; 40:1418–1425

49. Olds DL, Henderson CR, Cole R, Eckenrode J, Kitzman H, Luckey D, Pettitt L, Sidora K, Morris P, Powers J: Long-term effects of nurse home visitation on children's criminal and antisocial behavior: 15-year follow-up of a randomized controlled trial. JAMA 1998; 280:1238–1244

50. Cohen P, Cohen J, Kasen S, Velez CN, Hartmark C, Johnson J, Rojas M, Brook J, Streuning EL: An epidemiological study of disorders in late childhood and adolescence, I: age- and gender-specific prevalence. J Child Psychol Psychiatry 1993; 34:851–867

51. Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A: Prevalence and development of psychiatric disorders in childhood and adolescence. Arch Gen Psychiatry 2003; 60:837–844

52. Lahey BB, Loeber R, Quay HC, Applegate B, Shaffer D, Waldman I, Hart EL, McBurnett K, Frick PJ, Jensen P, Dulcan M, Canino G, Bird H: Validity of DSM-IV subtypes of conduct disorder based on age of onset. J Am Acad Child Adolesc Psychiatry 1998; 37:435–442

53. Egger HL, Angold A: The Preschool Age Psychiatric Assessment (PAPA): a structured parent interview for diagnosing psychiatric disorders in preschool children, in Handbook of Infant and Toddler Mental Health Assessment. Edited by DelCarmen-Wiggins R, Carter AS. New York, Oxford University Press, 2004, pp 223–243

54. Wakschlag L, Danis B: Assessment of disruptive behavior in young children: a clinical-developmental framework, in Handbook of Infant and Toddler Mental Health Assessment. Edited by DelCarmen-Wiggins R, Carter AS. New York, Oxford University Press, 2004, pp 421–440