Developmental Psychology
2007, Vol. 43, No. 1, 13–26

0012-1649/07/$12.00 DOI: 10.1037/0012-1649.43.1.13

Gender Differences in Physical Aggression: A Prospective Population-Based Survey of Children Before and After 2 Years of Age

Raymond H. Baillargeon
Université de Montréal

Mark Zoccolillo
McGill University

Kate Keenan
University of Chicago

Sylvana Côté and Daniel Pérusse
Université de Montréal

Hong-Xing Wu
Health Canada

Michel Boivin
Université Laval

Richard E. Tremblay
Université de Montréal

There has been much controversy over the past decades on the origins of gender differences in children's aggressive behavior. A widely held view is that gender differences emerge sometime after 2 years of age and increase in magnitude thereafter because of gender-differentiated socialization practices. The objective of this study was to test for (a) gender differences in the prevalence of physical aggression in the general population of 17-month-old children and (b) change in the magnitude of these differences between 17 and 29 months of age. Contrary to the differential socialization hypothesis, the results showed substantial gender differences in the prevalence of physical aggression at 17 months of age, with 5% of boys but only 1% of girls manifesting physically aggressive behaviors on a frequent basis. The results suggest that there is no change in the magnitude of these differences between 17 and 29 months of age.

Keywords: aggression, gender differences, epidemiology, longitudinal study, preschool

There has been much controversy over the past decades on the origins of gender differences in children's aggressive behavior (e.g., Cairns, 1979; Earls, 1987; Hyde, 1984; Lytton, 1990; Maccoby & Jacklin, 1980; Moffitt, Caspi, Rutter, & Silva, 2001; Tieger, 1980; Zahn-Waxler, 1993; Zoccolillo, 1993). Advocates of a social learning approach generally argue that gender differences do not exist early in life but only emerge later during the preschool period, at a time when gender-differentiated socialization pressures are believed to be well established. Keenan and Shaw (1997), for instance, argued that gender differences emerge because socializing agents, such as parents, selectively encourage traditional sex-type behaviors (e.g., shyness, fearfulness, and withdrawal in girls) and/or discourage non-sex-type behaviors (e.g., aggressive behavior in girls). In addition, according to a differential socialization hypothesis, girls are more responsive than boys to the socialization efforts of their parents and caregivers because they may present superior adaptive skills (e.g., language skills and empathic and prosocial behavior). As a result, the magnitude of gender differences in aggressive behavior is believed to increase over time. In contrast, those in favor of a biological basis argue that gender differences in aggressive behavior are present very early on in life before gender-differentiated socialization pressures could cause

Raymond H. Baillargeon, Department of Psychiatry, Université de Montréal, Montréal, Canada; Mark Zoccolillo, Department of Psychiatry, McGill University, Montréal, Canada; Kate Keenan, Department of Psychiatry, University of Chicago; Sylvana Côté, Department of Psychoeducation, Université de Montréal, Montréal, Canada; Daniel Pérusse, Department of Anthropology, Université de Montréal, Montréal, Canada; Hong-Xing Wu, Health Canada, Ottawa, Ontario, Canada; Michel Boivin, Department of Psychology, Université Laval, Québec City, Québec, Canada; and Richard E. Tremblay, Department of Psychology, Université de Montréal, Montréal, Canada.

Part of this article was presented at the 15th meeting of the International Society for Research on Aggression (Montréal, Québec, Canada). Raymond H. Baillargeon was supported by a Relève Médecine 2000 Award from the Faculty of Medicine, Université de Montréal. Preparation of this article was supported by funds from the Canadian Institutes of Health Research to Raymond H. Baillargeon and Mark Zoccolillo (Grant MOP-67090) and from the Social Sciences and Humanities Research Council of Canada to Richard E. Tremblay.

We thank Elisa Romano for her comments on an earlier version of this article. We also thank Géraldine Tanis for secretarial assistance, Jean-Michel Johnson for editorial assistance, and Suganthiny Jeyaganth for her help with some of the statistical analyses.

Correspondence concerning this article should be addressed to Raymond H. Baillargeon, Department of Psychiatry, Centre de Recherche de l'Hôpital Sainte-Justine, Université de Montréal, Room A-841, 3175 Côte-Sainte-Catherine, Montréal, Québec H3T 1C5, Canada. E-mail: raymond.baillargeon@recherche-ste-justine.qc.ca

them. Maccoby and Jacklin (1974), for instance, argued that boys and girls differ in their degree of preparedness toward aggression, with boys being more inclined to respond in aggressive ways to their environment. Other factors may account for the presence of gender differences early on in life, such as prenatal exposure to androgens, as well as pregnancy, labor, and delivery complications (Raine, 2002a, 2002b). At the heart of this controversy lie the questions of how early one can detect genuine gender differences in children's aggressive behavior and, once they have emerged, whether the magnitude of these differences is increasing over time.

The answer to this second question is especially critical to the origin of early gender differences in children's aggressive behavior. It may be that gender-differentiated socialization practices are being established and their effect on children's aggressive behavior is being felt earlier in development than previously thought (Fagot, 1984; Fagot & Hagan, 1985; Smetana, 1989). Similarly, toddlers may apply socially prescribed rules and standards to regulate their own aggressive behavior, for instance, girls shunning away from aggressive behavior to avoid self-criticism and maintain self-satisfaction and self-worth (Bussey & Bandura, 1999; Maccoby, 2002). Early gender differences in children's aggressive behavior could be culturally rather than biologically driven, but that would seem rather unlikely if their magnitude is not increasing over time. In this prospective population-based cohort study, we test not only for the presence of gender differences in physical aggression in children before 2 years of age but also for change in their magnitude during toddlerhood.

Investigating Gender Differences in Physical Aggression From a Developmental Epidemiology Perspective

How early can we detect genuine gender differences in children's physical aggression? From a developmental epidemiology perspective, this question refers to the prevalence of physical aggression in the general population of children before 2 years of age. It can be defined as the proportion of toddlers in the general population whose propensity to manifest physically aggressive behaviors (e.g., hitting, pushing) is much higher than that of other children of the same age. One reason for focusing on these toddlers is that physical aggression is not necessarily maladaptive in and of itself. However, when physically aggressive behaviors are manifested on a frequent/severe basis, it may well be so. In fact, many toddlers in the general population manifest these behaviors on an occasional basis, but there are relatively few who do so on a frequent basis (see below). Hence, although the latter may be considered physically aggressive, the former should not; therefore, only gender differences among the latter refer to atypical or pathological development. Focusing on physically aggressive toddlers may be especially important if a majority of them continue to manifest physically aggressive behaviors on a frequent basis after 2 years of age. There is some epidemiological evidence that frequent/severe specific behaviors, including forceful actions against peers present during the preschool years, can predict maladaptive behavior in adolescence and early adulthood (Caspi, Elder, & Bem, 1987; Stevenson & Goodman, 2001). Is it possible to distinguish between toddlers who manifest physically aggressive behaviors on a frequent basis and those who do so only occasionally or not at all? We investigated this issue and the one of gender differences in the prevalence of physical aggression before 2 years of age by using a logit-based latent class model (Hagenaars, 1993). Beside allowing for the identification of physically aggressive toddlers, another advantage of this latent class model is that it offers the potential to reconcile the inconsistent findings obtained from previous epidemiological studies that have found gender differences in some physically aggressive behaviors but not in others (see below). The gender paradox hypothesis stipulates that the gender least frequently affected by a childhood disorder is paradoxically the one more severely affected (Eme, 1992; Keenan, Loeber, & Green, 1999; Moffitt et al., 2001; Silverthorn & Frick, 1999). Accordingly, there may be fewer aggressive girls before 2 years of age, but they may be more likely than their male counterparts to manifest some physically aggressive behaviors on a frequent/severe basis, and, therefore, one would possibly not observe gender differences in these behaviors.

Does the magnitude of early gender differences in the prevalence of physical aggression increase over time? This question refers to the continuity and discontinuity in children's physical aggression status during toddlerhood. Some aggressive toddlers will stop manifesting physically aggressive behaviors on a frequent basis (i.e., remission); conversely, some nonaggressive toddlers will start doing so after 2 years of age (i.e., incidence). Are girls more likely to stop manifesting physically aggressive behaviors on a frequent basis, or are boys more likely to start doing so after 2 years of age, or both? Gender differences in the incidence and remission of physical aggression in toddlers, as well as the predictive accuracy of physical aggression before 2 years of age, were investigated with a time-specific latent variables (Markov) model. This model allows testing for change over time in the magnitude of gender differences in the prevalence of physical aggression. Evidence for change is present if there are gender differences in the prevalence of physical aggression after 2 years of age after controlling for such differences before 2 years of age. Another advantage of this model is that it allows investigating the mechanism behind such change (i.e., greater remission among girls, greater incidence among boys, or both).

Physically Aggressive Behaviors in Children Under 2 Years of Age

There is evidence from observational studies that the ability to selectively and strategically use physical force against one's peers[1] emerges around the time of the child's 1st birthday (for

[1] There is an emerging sense of ownership around this age whereby one's control over possessions is being acknowledged by peers (e.g., Bakeman & Brownlee, 1982). Breaking this social/moral contract may then result in struggles with companions, some involving the use of physical force in retaliation (e.g., Smetana, 1989). During the 2nd year of life, physical force also becomes a means for obtaining desirable ends (e.g., systematically varying forceful actions to tease peers or obtain objects). Further, similar to older children, 1-year olds most often respond negatively to "hard hits"—high-intensity hit—or any hit to the head (Brownlee & Bakeman, 1981), suggesting that such forceful actions have a negative valence early in life (for a different view, see Huesmann, 2002). Of course, the organization and the function of forceful actions will change over the

a review, see Hay, 2005; see also Tremblay, 2003), with the child's repertoire of physically aggressive behaviors diversifying and expanding during the 2nd year of life (e.g., Bronson, 1981; Restoin et al., 1985). For instance, in one of the earliest accounts of infants' social interactions, Bridges (1933) documented the occurrence of "aggressive attacks" involving biting, hair pulling, and hitting at around 14–15 months of age (see also Bronson, 1981; Brownlee & Bakeman, 1981; Bühler, 1931, 1935; Dunn & Munn, 1985; Eckerman, Whatley, & Kutz, 1975; Goodenough, 1931; Shirley, 1933). More recently, on the basis of mothers' retrospective reports, Tremblay et al. (1999) estimated the age of onset of some physically aggressive behaviors among children who manifested the behaviors in question at 17 months of age. The rate of onset was the steepest between 11 and 14–15 months of age and appeared to level off thereafter.

Epidemiological Surveys of Physically Aggressive Behaviors in Boys and Girls Under 2 Years of Age

Although narrative reviews of the literature suggest that gender differences in aggressive behavior are not present early in life (e.g., Keenan & Shaw, 1997; Loeber & Hay, 1997), a recent meta-analysis of observational studies suggests otherwise (Archer & Côté, 2005). Much of our knowledge on gender differences in aggressive behavior during infancy comes from small-scale studies that relied on nonprobability samples. Some studies have found that more boys than girls manifest aggressive behavior before 2 years of age (e.g., Fagot & Hagan, 1985; Hay, Castle, & Davies, 2000; Tremblay et al., 1999), whereas others have found no gender differences (e.g., Hay et al., 2000; Holmberg, 1980; Keenan & Shaw, 1994; Shaw, Keenan, & Vondra, 1994; Tremblay et al., 1999). To our knowledge, there are only three epidemiological surveys of physically aggressive behaviors in boys and girls under 2 years of age: the University of California Control Study (Macfarlane, Allen, & Honzik, 1954), the 1956 Child Health Survey (Heinstein, 1969; Hornberger, Bowman, Greenblatt, & Corsa, 1960), and the Québec Longitudinal Study of Child Development (Jetté, 2002; Jetté & Des Groseilliers, 2000; Plante, Courtemanche, & Des Groseilliers, 2002).[2] Each survey relied on a probability sample representative of a well-defined geographical unit.

The University of California Control Study is the first epidemiological study of behavioral and emotional problems in children under 2 years of age (Macfarlane et al., 1954). This longitudinal study was conducted by Jean W. Macfarlane and her colleagues on a representative sample of children born in Berkeley, California, in the late 1920s. Mothers' reports before children were 5 years of age were dichotomized on the basis of the presence or absence of a problem. Macfarlane et al. found that 59% of boys (i.e., 33 out of 56) and 43% of girls (i.e., 26 out of 60) were reported as having temper tantrums, some including biting, kicking, and striking at 21 months of age (Cohen's, 1988, effect size index, **w** = .16). Neither these gender differences nor those observed in this problem 15 months later when children reached 36 months of age were statistically significant.

The 1956 Child Health Survey was conducted over a 10-week period beginning June 4, 1956, on a representative sample of the noninstitutionalized population of children under 6 years of age from the State of California (Heinstein, 1969; Hornberger et al., 1960). Mothers were asked to rate the intensity of behavioral and emotional problems as mild, intermediate, or severe. The authors found that 7% of 12- to 17-month-old boys (i.e., 4 out of 54) and 16% of girls (i.e., 5 out of 32) were reported as fighting with other children, but none severely (**w** = .13). Similarly, 18% of 18- to 23-month-old boys (i.e., 7 out of 38) and 36% of girls (i.e., 21 out of 59) were reported as fighting with other children, but only 1 boy and 1 girl (i.e., 2 out of 97) severely (**w** = .20). The authors did not report on the statistical significance of gender differences for specific age groups. However, if we assume a random sample (it was actually a single-stage, stratified cluster sample), it can be determined that these gender differences in fighting are not statistically significant.

The Québec Longitudinal Study of Child Development (QLSCD) is following a representative birth cohort of children born in 1997–1998 to mothers living in the Canadian province of Québec. Mothers were asked to rate the frequency of five physically aggressive behaviors (i.e., fights, attacks, kicks, bites, and hits) as never, sometimes, or often. At 17 months of age, between 4.5% and 24.1% of children manifested physically aggressive behaviors on an occasional basis; in contrast, less than 5% of children manifested these behaviors on a frequent basis (Baillargeon, 2002; Baillargeon et al., 2002, in press; Baillargeon & Zoccolillo, 2003). Boys were more likely than girls to kick (2.4% vs. 1.1%; **w** = .08), bite (4.9% vs. 3.1%; **w** = .07), and hit (0.9% vs. 0.3%; **w** = .09) other children on a frequent basis; in addition, the magnitude of gender differences in these behaviors—with the exception of fighting and attacking—increased between 17 and 29 months of age (Baillargeon, 2002; Baillargeon et al., 2002, in press; Baillargeon & Zoccolillo, 2003).

Objectives

Overall, the above findings show that physically aggressive behaviors are not typical of toddlers in the general population, at least when manifested on a frequent basis; in addition, the findings

first few years of life (e.g., Maudry & Nekula, 1939). The focus of this study, however, is not to describe the conditions that bring forth these forceful actions and the behavioral context in which they appear, but rather to study mothers' reports of specific physically aggressive behaviors as they co-occur within the same toddler over time.

[2] There are at least six other large-scale studies of behavior in children under 2 years of age. Four epidemiological studies did not report on aggressive behavior (Jenkins, Bax, & Hart, 1980; Janson, 2003a, 2003b; Mathiesen & Sanson, 2000) or relied only on parents' concerns (Blumberg, Halfon, & Olson, 2004). Two recent studies from the United States reported on a heterogeneous scale including aggression but also other types of disruptive behaviors (Achenbach & Rescorla, 2000; Briggs-Gowan, Carter, Skuban, & Horwitz, 2001). Moreover, they excluded unhealthy children at birth and used rather broad age groupings when reporting their results on aggressive behavior items (Achenbach & Rescorla, 2000) or scales (Carter, Briggs-Gowan, Jones, & Little, 2003), lumping together children under 2 years of age with children over 2 years of age. Other population surveys of children under 2 years of age have not generally included detailed assessment of children's aggressive behaviors (e.g., Anderson, 1936/1972; Bone & Meltzer, 1989; Golding, Pembrey, Jones, & ALSPAC Study Team, 2001; Kasmini, Kyaw, Krishnaswamy, Ramli, & Hassan, 1993; U.S. National Health Survey, 1959).

also showed that gender differences are already present in at least some physically aggressive behaviors before 2 years of age. Many important questions remain unanswered, however. Is there a greater number of boys than girls in the general population who manifest physically aggressive behaviors on a frequent basis prior to age 2? Are girls more likely than boys to stop manifesting physically aggressive behaviors on a frequent basis, or are girls less likely than boys to start doing so during toddlerhood, or both? The first objective of this study is to test for gender differences in the prevalence of physical aggression in the general population of children at 17 months of age by using a logit-based latent class model. The second objective of this study is to test for change in the magnitude of these differences between 17 and 29 months of age by using a time-specific latent variables (Markov) model.

Method

Sample

The QLSCD is conducted by Santé Québec, a division of the Institut de la Statistique du Québec (ISQ). It is following, on an annual basis, a representative birth cohort of children born between October 1997 and July 1998 to mothers living in the province of Québec, Canada. At Wave 1 in 1998, the children were between 59 and 64 weeks of gestational age—gestational age was defined as the sum of the duration of pregnancy and the chronological age of the baby. (At Wave 2 and Wave 3, children were approximately 17 and 29 months old, respectively.) Children living in Northern Québec, Cree territory, Inuit territory, and Indian reserves and those for whom the duration of the pregnancy could not be determined were excluded from the target population. The target population for Wave 1 was representative of approximately 96.6% of the Québec population of newborns. The size of the target population was slightly smaller for subsequent waves mainly because children from families who arrived in Québec after July 1998 were not recruited. Note that 79,724 and 75,674 children were born in Québec in 1997 and 1998, respectively (Institut de la Statistique du Québec, 2002).

Infants were selected from the 1997–1998 Master Birth Register of the Ministry of Health and Social Services, which contains records of all birth certificates by calendar year. Access to this information was obtained with the prior approval of the Québec Access to Information Commission. Some infants were excluded from the first wave of data collection because they were not yet listed in the register at the time of selection. Infants born either before 24 weeks or after 42 weeks of gestation were also excluded. In addition, there was a slight undercoverage of infants whose gestational age was between 24 and 36 weeks. Finally, infants were excluded if their gender could not be determined from the register. In all, 2,940 infants were selected. They were representative of 94.5% of the target population.

Instrument

The Interviewer Completed Computerized Questionnaire (ICCQ) is the main questionnaire of the QLSCD. It was administered during a face-to-face interview conducted in the child's home with the person most knowledgeable about the child, who in over 99% of the cases was the child's biological mother. From Wave 2 onward, when children were about 17 months old, the ICCQ collected information on a wide variety of behaviors, some of which are typically associated with later child psychiatric disorders like oppositional-defiant disorder, attention-deficit/hyperactivity disorder, and conduct disorder. Five behavior items that necessarily involve physical aggression were considered in this study: (a) gets into many fights, (b) physically attacks people, (c) kicks other children, (d) bites other children, and (e) hits other children. Two behavior items (i.e., fights and attacks) came from the Child Behavior Checklist for Ages 2–3 (Achenbach, Edelbrock, & Howell, 1987) that was originally modeled on instruments for rating preschool- and school-age children by parents and teachers among others (Achenbach & McConaughy, 1997). The other three behavior items (i.e., kicks, bites, and hits) were adapted from the Preschool Behavior Questionnaire (Behar & Stringfield, 1974; Fowler & Park, 1979), an adaptation of the Children's Behaviour Questionnaire (Rutter, 1967). Each behavior item was rated by the child's biological mother on a 3-point Likert scale: 1 = *doesn't apply or never,* 2 = *occasional behavior or sometimes,* and 3 = *frequent behavior or often.* Mothers are generally considered valid informants of children's behavior, especially during the preschool years, because they are likely to be familiar with their child's behavior across different settings (e.g., Carter et al., 2003; Earls, 1980).

Note that we collapsed the data for the kick, bite, and hit behavior items into a single composite behavior item by using the highest rating category on any one of these three behavior items. In other words, children who received a rating of 3 (i.e., often) on any of the three behavior items in question were given a 3 on the composite behavior item; among the remaining children, those who received a rating of 2 (i.e., sometimes) on any of the three items were given a 2 on the composite item; otherwise, they were given a rating of 1 (i.e., never). This was done in an attempt to avoid large sparse multidimensional tables that may jeopardize the asymptotic suitability of the chi-square distribution for the Pearson and likelihood-ratio chi-square statistics (see below). This is especially appropriate because the third rating category (i.e., often) was generally endorsed by relatively few mothers. Moreover, this procedure did not yield a statistically significant reduction in the amount of variation present in the longitudinal data on behavior items as assessed by the method described by Goodman (1981). Baillargeon, Tremblay, and Willms (1999) used data from the National Longitudinal Survey of Children and Youth and found that the reliability of a measure made up of three very similar behavior items for 2-year-old children was good (i.e., fights; kicks, bites, hits other children; and reacts with anger and fighting when accidentally hurt)—that is, .87 and .84 for boys and girls, respectively. In addition, very similar measures were used in previous longitudinal studies to assess, for instance, trajectories of physical aggression during early childhood as well as correlates of high-level physical aggression trajectory (Tremblay et al., 2004; see also Broidy et al., 2003; Cairns, Cairns, Neckerman, Ferguson, & Gariépy, 1989).

Among the 2,940 infants who made up the first wave sample, 123 infants did not take part in the subsequent waves. Among the 2,817 who took part in the longitudinal survey, 2,120 were ICCQ respondents (i.e., 8 ineligible; 689 nonrespondents), yielding an ICCQ cross-sectional weighted response rate of 73.5% (i.e., weighted ratio of number of respondents over total number of eligible respondents). The sociodemographic characteristics of these infants' families are presented in Table 1 (see also Desrosiers, 2000). Note that 41.5% of these infants were firstborn, and 79.8% of families were still intact when the child was 5 months old. The annual income of 33.3% of families was below 30,000 Canadian dollars, and welfare represented the principal source of income for 11.7% of households. Note also that 15.7% of parents were immigrants; 3.1% of mothers but only 0.3% of fathers were under 20 years of age; and 18.0% of mothers and 17.6% of fathers did not have a high school diploma. Infants were from various cultural and linguistic backgrounds. In all, 18% of infants did not belong to one of the majority ethnocultural groups (i.e., Canadian-native-born, French, British, Irish, or Scottish). Other declared ethnocultural origins were African/Haitian, aboriginal (Amerindian), Spanish speaking (of the Americas), and Arab. Infants were also from various linguistic backgrounds, with 8% living in a household in which the language most often spoken was neither French—the only language spoken at home in 75% of households—nor English. Among the 2,120 Wave 1 ICCQ respondents, 1,985 were included in Waves 2 and 3 (i.e., 19 ineligible; 116 nonrespondents), yielding an ICCQ longitudinal weighted response rate of 70.5%. In fact, the attrition rate was very low, with only 3.5% of the Wave 1 ICCQ respondents still eligible for the two subsequent waves not responding to

Table 1
Sociodemographic Characteristics of the Infants' Household at Wave 1 of the Québec Longitudinal Study of Child Development

Variable	Overall %	Breakdown by parent: Mother (%)	Breakdown by parent: Father (%)
No. of brothers and sisters (n = 2,120)			
None	41.5		
1	40.2		
2	12.2		
3 or more	6.2		
Family composition (n = 2,112)			
Two biological parents	79.8		
Step family	10.8		
Single-parent family	9.3		
Ethnocultural origin of the mother (n = 2,118)			
Nonimmigrant	84.3		
Immigrant of European origin	3.3		
Immigrant of non-European origin	12.4		
Household annual income, in Canadian dollars (n = 2,077)			
<30,000	33.3		
30,000–59,999	39.5		
≥60,000	27.2		
Principal source of family income (n = 2,092)			
Salary, wage, or self-employment	83.5		
Welfare	11.7		
Unemployment insurance	1.8		
Other	3.0		
Educational attainment[a]			
No high school diploma		18.0	17.6
High school diploma		11.4	12.6
Postsecondary studies		17.7	16.7
Vocational/technical school diploma		10.8	11.5
College diploma		17.5	16.7
University degree		24.7	24.8
Parents' age (in years)[b]			
<20		3.1	0.3
20–24		19.8	8.0
25–29		30.4	26.7
30–34		32.5	34.2
35–39		11.8	22.1
≥40		2.5	8.7

[a] Mother, n = 2,116; Father, n = 1,905. [b] Mother, n = 2,119; Father, n = 1,924.

Wave 2 ICCQ. Similarly, only 5.8% of the Wave 1 ICCQ respondents still eligible for the two subsequent waves did not respond to the Wave 3 ICCQ. In this study, sampling weights designed by the ISQ to reduce cross-sectional and longitudinal overall nonresponse bias as well as undercoverage of the target population were used to weight the data.

Statistical Method

Testing for gender differences in prevalence of physical aggression at 17 months of age. A child's physical aggression status at 17 months of age was inferred from his or her data on the behavior items through a logit-based latent class model (Hagenaars & McCutcheon, 2002; Rost & Langeheine, 1997; von Eye & Clogg, 1994). We considered a latent class model with one latent variable made up of three latent classes: a low-, medium-, and high-aggressive latent class. It was assumed that the behavior items elicited ratings in the same category for a given child. Members of the low-aggressive latent class would tend not to manifest physically aggressive behaviors, whereas members of the medium- and high-aggressive latent classes would tend to manifest these behaviors on an occasional and frequent basis, respectively. Under this model, the physical aggression latent variable explained the interrelationships among the behavior items, with each child being in one, and only one, latent class. We estimated the conditional probability of a randomly selected boy and girl in the population manifesting a particular behavior item never, sometimes, or often, given his or her latent class membership. Further, we estimated the probability of a randomly selected boy and girl in the population belonging to the low-, medium-, and high-aggressive latent class. Further, we also estimated the boy–girl ratio of the odds of belonging to the high- rather than medium-aggressive latent class; and, similarly, we estimated the boy–girl ratio of the odds of belonging to the medium- rather than low-aggressive latent class.

Testing for change in magnitude of gender differences in prevalence of physical aggression between 17 and 29 months of age. The continuity and discontinuity in a child's latent physical aggression status from 17 to 29 months of age were inferred from his or her longitudinal data on the behavior items through a time-specific latent variables (Markov) model (Baillargeon et al., 2004; Langeheine & van de Pol, 1994; van de Pol & Langeheine, 1990). We considered a model with two time-specific latent variables each made up of a low-, medium-, and high-aggressive latent

class. Note that the 17-month-old latent physical aggression variable and gender were independent (i.e., explanatory) variables with fixed two-dimensional margins (i.e., there was no constraint on their association), whereas the 29-month-old latent physical aggression variable was a dependent (i.e., response) variable. Under this model, each latent variable explained the interrelationships among the behavior items at a given time; hence, it was assumed to have no effect on the behavior items at another time. For each latent variable, we estimated the conditional probability of a randomly selected boy and girl in the population manifesting a particular behavior item never, sometimes, or often given his or her latent class membership. Further, we estimated the conditional probability of a randomly selected boy and girl in the population belonging to the low-, medium- or high-aggressive latent class at 29 months of age, given he or she was low-aggressive, medium-aggressive, and high-aggressive, respectively, 1 year earlier. Further, we estimated simultaneously for boys and girls the association between a child's latent physical aggression status at 17 and 29 months of age. Note that we excluded children who did not have complete data on all behavior items at 17 and 29 months of age; only 3 out of 1,985 ICCQ longitudinal respondents were excluded for this reason.

Because this model took into account gender differences in the prevalence of physical aggression at 17 months of age, the absence of association between gender and the 29-month-old latent physical aggression variable (beyond that expected by chance alone) would suggest that there were no gender differences in the incidence and remission of physical aggression between 17 and 29 months of age and, therefore, no change in the magnitude of gender differences in the prevalence of physical aggression during this period.

Maximum likelihood estimation and assessing model goodness-of-fit. Maximum likelihood parameter estimates of the different statistical models considered in this study were obtained using the expectation maximization (EM) algorithm from lEM (Vermunt, 1997), a computer program for the analysis of categorical data. The EM algorithm was run at least 1,000 times with different starting values. Each time the iterations were stopped when a convergence criterion was reached (i.e., the minimum increase in the log-likelihood function between subsequent iterations was set at .0000000000000001) or when a certain number of iterations was performed (i.e., 10,000). This was done to ensure that the maximum likelihood estimates represented a global rather than local maximum (McCutcheon, 1987). The fit of a particular statistical model to the QLSCD data was assessed using the Pearson chi-square (X^2), the likelihood-ratio chi-square (L^2), and the Cressie–Read (CR) statistics. The X^2, L^2, and CR statistics have a large sample chi-square distribution under certain conditions (Clogg, 1979). However, discrepancies between the X^2 and the L^2 statistics may be due to a relatively large number of cells with zero or near-zero observed frequencies—these cells receive zero weight in the summary of fit that the L^2 provides. The CR statistic represents a middle ground with a weight that is neither 1 (as in the X^2) nor 0 (as in the L^2). Cressie and Read (1984) recommended a weight of 2/3 for sparse data. Whenever the expected frequencies are close to the observed frequencies, the X^2/L^2/CR value will be small relative to the degrees of freedom (i.e., the number of nonredundant observed cell frequencies minus the number of independent parameters to be estimated), and the model being examined can be said to provide an adequate fit to the data. Two hierarchically related models can be compared using the L^2 because it can be partitioned exactly, which is not the case for the X^2 or CR (Fienberg, 1980). Two models are hierarchically related if one includes all the parameters of the other plus some additional ones. Similarly, two models can be compared using the Bayesian information criterion (BIC): L^2 – degrees of freedom * natural logarithm (N). The model with the smallest BIC value is to be preferred.

To take into account the QLSCD's design effect (i.e., the variance of the parameter estimates obtained from the QLSCD is likely to be underestimated by approximately 30% because of its stratified rather than simple random sampling design), which increased the risk of falsely rejecting the null hypothesis, we have used a conservative alpha level (i.e., $\alpha = .01$).

Results

Prevalence of Physical Aggression at 17 Months of Age

Before we could test for gender differences in the prevalence of physical aggression at 17 months of age, we needed to determine whether a three-class model was appropriate for the 17-month-old data on behavior items. Table 2 presents the goodness-of-fit test statistics associated with the different statistical models considered in this study. We considered a model with one latent physical aggression variable made up of three mutually exclusive and exhaustive latent classes. The value of the L^2 associated with this model was 21.67, with 13 degrees of freedom ($p = .06$), suggesting that this model was appropriate. Furthermore, there were no standardized residuals greater than 2.58 in absolute value associated with this model. Using the one-class model as a benchmark, we estimated that this model accounted for 96% (1 – [21.67/537.43]) of the observed variation in the 17-month-old data on behavior items. Note that the conventional symptom loading approach (Glidewell, Mensh, & Gildea, 1957) did not seem adequate for the 17-month-old data on behavior items. In fact, the L^2 associated with a Partial Credit model (Masters, 1982), which assumed a continuous distribution of physical aggression and that the behavior items were interval data with equal discriminating power, was 83.93, with 38 degrees of freedom ($p = .000$). Hence, on all accounts, this three-class model seemed to provide an excellent fit to the 17-month-old data on behavior items.

Table 3 presents the estimates of the conditional probability of a randomly selected boy and girl in the population manifesting a particular behavior item never, sometimes, or often, given his or her latent class membership. As expected, the three-class model allowed us to distinguish between children who manifested physically aggressive behaviors on a frequent basis and those who did so only occasionally. For instance, consider the odds of fighting often rather than sometimes for girls. The odds were 2.72 (0.73/0.27) for high-aggressive girls; comparatively, the odds were only 0.06 (0.04/0.64) for medium-aggressive girls (see Table 3). Hence, the odds of fighting often rather than sometimes were 46.4 (2.72/0.06) times higher for high- than for medium-aggressive girls. In addition, as expected, the three-class model allowed us to distinguish between children who manifested physically aggressive behaviors on an occasional basis and those who tended not to manifest these behaviors. For instance, consider the odds of fighting sometimes rather than never for girls. The odds were 1.96 (0.64/0.33) and 0.06 (0.06/0.94) for medium- and low-aggressive girls, respectively (see Table 3). Hence, the odds of fighting sometimes rather than never were 32.6 (1.96/0.06) times higher for medium- than for low-aggressive girls.

Gender Differences in the Prevalence of Physical Aggression at 17 Months of Age

Table 3 presents the estimates of the probability of a randomly selected boy and girl in the population belonging to the low-, medium-, and high-aggressive latent class. As expected, there was a substantial number of medium-aggressive but a much smaller number of high-aggressive children at 17 months of age. Whereas 35% of boys in the population were medium-aggressive, only 5% were high-aggressive. Comparatively, there were 18% of medium-

Table 2
Goodness-of-Fit Test Statistics Associated With Different Statistical Models

Model	X^2	*p*	L^2	*p*	CR	*p*	*df*	BIC
Prevalence of physical aggression at 17 months of age								
Logit-based latent class model								
One-class	2,229.83	.000	537.43	.000	1,057.57	.000	40	233.78
Two-class	159.45	.000	96.84	.000	123.83	.000	26	−100.54
Three-class[a]	17.80	.17	21.67	.06	18.45	.14	13	−77.02
Three-class with no gender differences in the prevalence of physical aggression	26.47	.02	30.61	.01	27.23	.02	14	−75.67
Three-class with no between-gender variation in the 17-month-old data on behavior items	63.31	.000	69.98	.000	64.03	.000	32	−172.94
Continuity and discontinuity in children's latent physical aggression status between 17 and 29 months of age								
Time-specific latent variables (Markov) model								
Unconstrained	1,695.90	.000	768.37	1.0	1,086.61	1.0	1,368	−9,614.35
No change in the magnitude of gender differences in the prevalence of physical aggression over time	1,718.47	.000	781.85	1.0	1,105.62	1.0	1,374	−9,646.42
Equal likelihood of stopping vs. starting manifesting physically aggressive behaviors on a frequent basis over time	1,697.65	.000	789.47	1.0	1,108.57	1.0	1,376	−9,653.97
Equal stability over time of the low-, medium-, and high-aggressive statuses	1,680.88	.000	784.98	1.0	1,098.24	1.0	1,376	−9,658.47

Note. X^2 and L^2 are the Pearson and likelihood-ratio chi-square statistic, respectively; CR is the Cressie–Read statistic. BIC = Bayesian information criterion.
[a] The boy–girl ratio of the odds of belonging to the high- rather than medium-aggressive latent class was set equal to the boy–girl ratio of the odds of belonging to the medium- rather than low-aggressive latent class.

aggressive and 1% of high-aggressive girls in the population. Moreover, these results suggest a strong gender effect on the prevalence of physical aggression at 17 months of age. In fact, the odds of belonging to the high- rather than medium-aggressive latent class were 2.62 times higher for boys than for girls (99% confidence interval = 1.31–5.24); and, similarly, the odds of belonging to the medium- rather than low-aggressive latent class were 2.62 times higher for boys than for girls.

How much of the observed variation in the 17-month-old data on behavior items was due to gender differences in the prevalence of physical aggression? To answer this question, we considered a restricted three-class model that assumed there was no association between the latent physical aggression variable and gender (see Table 2, top). Under this model, there were no gender differences in the prevalence of physical aggression beyond those expected in the propensity to manifest a particular behavior item for a given

Table 3
Parameter Estimates Under the Three-Class Model for the 17-Month-Old Data on Behavior Items

	Low-aggressive		Medium-aggressive		High-aggressive	
Rating	Boy	Girl	Boy	Girl	Boy	Girl
Probability of belonging to a particular latent class (i.e., prevalence)						
	.604 (.06)	.812 (.03)	.346 (.06)	.178 (.03)	.050 (.02)	.01 (.004)
Conditional probability of fighting given membership in a particular latent class						
Never	.98 (.02)	.94 (.02)	.61 (.07)	.33 (.08)	.15 (.09)	—
Sometimes	.02 (.02)	.06 (.02)	.39 (.06)	.64 (.08)	.50 (.10)	.27 (.22)
Often	.00 (.002)	—	.004 (.01)	.04 (.02)	.34 (.12)	.73 (.22)
Conditional probability of attacking given membership in a particular latent class						
Never	.96 (.02)	.94 (.01)	.69 (.05)	.46 (.07)	.20 (.14)	.18 (.22)
Sometimes	.04 (.02)	.06 (.01)	.31 (.04)	.52 (.07)	.32 (.09)	—
Often	.006 (.004)	.004 (.004)	.002 (.02)	.02 (.02)	.49 (.14)	.82 (.22)
Conditional probability of kicking, biting, and hitting given membership in a particular latent class						
Never	.83 (.04)	.75 (.02)	.23 (.07)	.29 (.06)	.11 (.07)	.60 (.18)
Sometimes	.17 (.04)	.23 (.02)	.64 (.06)	.60 (.06)	.20 (.13)	.15 (.14)
Often	—	.02 (.01)	.13 (.04)	.11 (.03)	.68 (.15)	.25 (.16)

Note. Standard errors appear in parentheses. A dash indicates that no estimates were obtained.

latent physical aggression status. This is why these differences (in the prevalence of physical aggression) are sometimes referred to as "true" group differences (Thissen, Steinberg, & Gerrard, 1986). Using the unrestricted one-class model as a benchmark, we estimated that only 1.7% ([30.61 – 21.67]/537.43) of the observed variation in the 17-month-old data on behavior items was due to gender differences in the prevalence of physical aggression (**w =** .24, between a small and medium effect size). This is not to say, however, that the remaining 98% of the observed variation was due to within-gender variation. To evaluate the relative importance of within- versus between-gender variation, we considered a restricted three-class model, which assumed that there was no between-gender variation in the 17-month-old data on behavior items (see Table 2, top). Under this model, there were neither gender differences in the prevalence of physical aggression nor gender differences in the propensity to manifest a particular behavior item for a given latent physical aggression status. Using the one-class model as a benchmark, we estimated that 9% ([69.98 – 21.67]/537.43) of the observed variation in the 17-month-old data on behavior items was due to between-gender variation. Hence, although gender differences in the prevalence of physical aggression at 17 months of age were substantial, most of the observed variation in the 17-month-old data on behavior items was due to within-gender variation.

Continuity and Discontinuity in Children's Latent Physical Aggression Status From 17 to 29 Months of Age

Before we could test for change in the magnitude of gender differences in the prevalence of physical aggression between 17 and 29 months of age, we needed to determine whether a time-specific latent variables (Markov) model was appropriate for the longitudinal data on behavior items. The value of the L^2 associated with this model was 768.37, with 1,368 degrees of freedom (p = 1.0). Note that although the X^2 statistic suggested that this model did not fit the data, the CR statistic yielded the same conclusion as the L^2 (see Table 2, bottom). Overall, a time-specific latent variables (Markov) model seemed to provide an acceptable fit to the longitudinal data on behavior items.

Change in the magnitude of gender differences in the prevalence of physical aggression between 17 and 29 months of age. We considered a restricted time-specific latent variables (Markov) model, which assumed that there was no association between gender and the 29-month-old latent physical aggression variable beyond that expected by chance alone after controlling for gender differences in the prevalence of physical aggression at 17 months of age (see Table 2, bottom). Note that this model also implied that there were no gender differences in the association between the 17- and 29-month-old latent physical aggression variables. Under this model, high-aggressive 17-month-old boys were just as likely as their female counterparts to stop manifesting physically aggressive behaviors on a frequent basis 1 year later; similarly, low-aggressive 17-month-old girls were just as likely as their male counterparts to start manifesting physically aggressive behaviors 1 year later. This model represented an increase in L^2 of 13.48, with a corresponding increase of 6 degrees of freedom from the previous model (L^2 = 781.85 – 768.37 = 13.48; df = 1,374 – 1,368 = 6; p = .04). These results suggest that there was neither an increase nor a decrease in the magnitude of gender differences in the prevalence of physical aggression between 17 and 29 months of age.

Table 4 presents the estimates of the conditional probability of a randomly selected 29-month-old child in the population belonging to the low-, medium-, or high-aggressive latent class given his or her latent physical aggression status at 17 months of age. On the one hand, about half (i.e., 51.1%) of 17-month-old high-aggressive children continued to manifest physically aggressive behaviors on a frequent basis 1 year later. On the other hand, a majority (i.e., 82.8%) of 17-month-old low-aggressive children continued not manifesting physically aggressive behaviors 1 year later.

Discontinuity in children's latent physical aggression status from 17 to 29 months of age. Between 17 and 29 months of age, many aggressive toddlers did stop manifesting physically aggressive behaviors on a frequent basis, and many nonaggressive toddlers did start doing so. Which was the most likely transition? To answer this question, we considered a restricted time-specific latent variables (Markov) model wherein (a) the likelihood of changing from low- to high-aggressive was equal to that of changing from high- to low-aggressive, and (b) the likelihood of changing from medium- to high-aggressive was equal to that of changing from high- to medium-aggressive (see Table 2, bottom). This conditional symmetry model represented an increase in L^2 of only 7.62, with a corresponding increase of 2 degrees of freedom from the preferred time-specific latent variables (Markov) model (L^2 = 789.47 – 781.85 = 7.62; df = 1,376 – 1,374 = 2; p = .02); in addition, this model had a smaller BIC value. These results suggest that nonaggressive toddlers were just as likely to start manifesting physically aggressive behaviors on a frequent basis as aggressive toddlers were to stop doing so between 17 and 29 months of age. Moreover, because there were fewer high- than low- and medium-aggressive children at 17 months of age, these results suggest that there were more children "drifting into" than "drifting out of" manifesting physically aggressive behaviors on a frequent basis between 17 and 29 months of age. In fact, under this conditional symmetry model, 2.3% of boys drifted into, but only 0.5% drifted out. Similarly, 2.1% of girls drifted into, but only 0.06% drifted out. Hence, there seemed to be a tendency for the prevalence of physical aggression to increase during this period.

Continuity in children's latent physical aggression status from 17 to 29 months of age. Was the stability of a child's physical aggression status between 17 and 29 months of age related to its severity? To answer this question, we considered a restricted time-specific latent variables (Markov) model wherein the likelihood of staying low-aggressive was equal to that of staying medium-aggressive, which in turn was equal to that of staying high-aggressive (see Table 2, bottom). This equal stability model represented an increase in L^2 of 3.13, with a corresponding increase of 2 degrees of freedom from the preferred time-specific latent variables (Markov) model (L^2 = 784.98 – 781.85 = 3.13; df = 1,376 – 1,374 = 2; p = .21), suggesting that the 1-year stability of a 17-month-old child's physical aggression status was not related to its severity. In fact, under this equal stability model, 80% of 17-month-old children were estimated to be in the same physical aggression status 1 year later. Hence, these results suggest that the predictive value of a 17-month-old child's latent physical aggression status was excellent, at least as far as predicting his or her status 1 year later was concerned.

Table 4
Estimates of Conditional Probability of a Randomly Selected 29-Month-Old Child Belonging to the Low-, Medium-, or High-Aggressive Latent Class Given His or Her Latent Physical Aggression Status at 17 Months of Age

Latent physical aggression status at 17 months of age	Latent physical aggression status at 29 months of age		
	Low-aggressive	Medium-aggressive	High-aggressive
Low-aggressive	.828 (.025) [.514; .594]	.171 (.025) [.106; .123]	.0005 (.0003) [.0003; .0003]
Medium-aggressive	.190 (.042) [.059; .052]	.768 (.044) [.237; .210]	.042 (.015) [.013; .011]
High-aggressive	.006 (.004) [.0004; .0001]	.483 (.096) [.035; .005]	.511 (.098) [.037; .005]

Note. These estimates were obtained from a restricted time-specific latent variables (Markov) model, wherein there is a uniform association between the two latent variables. This model represented an increase in L^2 of only 2.58, with a corresponding increase of 3 degrees of freedom from the preferred model ($L^2 = 784.42 - 781.85 = 2.58$; $df = 1{,}377 - 1{,}374 = 3$; $p = .46$), suggesting that a single odds ratio was sufficient to describe the association between the two latent variables. The value of this odds ratio was estimated at 19.5 (99% confidence interval = 9.92–38.34). Standard errors of estimates appear in parentheses, and unconditional probability estimates appear in brackets with the boys' estimates presented first.

Discussion

There is widespread consensus that school-age boys manifest more disruptive behaviors such as physical aggression than their female counterparts (for a review, see Lahey, Miller, Gordon, & Riley, 1999). What is the origin of these gender differences? A widely held view is that gender differences emerge sometime after 2 years of age because of gender-differentiated socialization pressures, whereby caregivers discourage aggressive behavior in girls for whom it is judged especially undesirable (Keenan & Shaw, 1997). However, there is considerable controversy as to whether such gender-differentiated socialization practices occur—and to what extent—relative to aggressive behavior (e.g., Lytton & Romney, 1991).

Origin of Gender Differences in Aggressive Behavior Before 2 Years of Age

Contrary to the differential socialization hypothesis, our results show that gender differences in the prevalence of physical aggression are already present in the 2nd year of life. Of course, these differences may at least, in part, be socially driven, with the effect of gender-differentiated socialization practices on early aggressive behavior being felt before the child's 2nd birthday. Our study did not entail any direct measure of such practices, and therefore that possibility cannot be ruled out. However, a unidirectional effect of gender-differentiated practices or of gender-stereotypic self-regulatory controls on early aggressive behavior seems unlikely in light of our longitudinal results, which suggest that girls are not more likely than boys to stop manifesting physically aggressive behavior between 17 and 29 months of age. These results are consistent with two observational studies. Hay et al. (2000) found no change in the magnitude of gender differences for both the rate of peer-to-peer use of force and mothers' ratings of their child's aggressiveness between 18 and 24 months of age. Similarly, Shaw et al. (1994) found no change in the magnitude of gender differences in a global rating of socially and nonsocially appropriate aggression between 18 and 24 months of age. It is possible, however, that gender-differentiated socialization practices had an effect on the continuity and discontinuity in children's latent physical aggression status from 17 to 29 months of age, but that it was canceled because of girls maturing faster than boys in the general context of an increase in the prevalence of physical aggression over time. In contrast to gender-differentiated practices, such differential in maturational tempo would make girls less likely than boys to stop manifesting physically aggressive behaviors during this period. Overall, it seems unlikely that gender-differentiated socialization practices or gender-stereotypic self-regulatory controls could be responsible for the emergence of gender differences in the prevalence of physical aggression prior to age 2.

How Substantial Are Early Gender Differences in the Prevalence of Physical Aggression?

Gender differences in the prevalence of physical aggression account for only a small percentage of the observed variation in physically aggressive behaviors among 17-month-old children, but our results show that these differences are substantial. In fact, for every physically aggressive girl, there were 5 boys who manifested physically aggressive behaviors on a frequent basis at 17 months of age. Incidentally, this boy–girl ratio is similar to ratios obtained in epidemiological studies of conduct disorder among 4- to 18-year-old children and youths (Lahey et al., 1999; see also Links, 1983). These boy–girl ratios have been said to exaggerate gender differences (Rossi, 1983). One way to appreciate how substantial gender differences in the prevalence of physical aggression really are is to consider the population attributable fraction (Fleiss, Levin, & Paik, 2003): the fraction of the prevalence of physical aggression in the population that can be uniquely attributable to being a boy. In fact, if boys were no more likely than girls to manifest physically aggressive behaviors on a frequent basis, the prevalence of physical aggression in the general population of

children at 17 months of age would be reduced by 67%.[3] Another way is to consider the attributable fraction among boys only. In fact, 80% (i.e., [.05 – .01]/.05) of 17-month-old physically aggressive boys in the general population might not have been expected to manifest physically aggressive behaviors on a frequent basis if it were not for the excess of prevalence among boys.

Gender Paradox Hypothesis

Typically, measures of physical aggression show larger gender differences than any other kind of aggressive behavior in children (Hyde, 1984; Knight, Fabes, & Higgins, 1996). Most of the between-gender variation in the 17-month-old data on behavior items (i.e., 9% of the observed variation) was due to gender differences in the propensity to manifest a particular behavior item for a given latent physical aggression status. Consistent with the gender paradox hypothesis, our results show that high-aggressive 17-month-old girls seem more likely than their male counterparts to fight and attack other children on a frequent basis (see Table 3). An alternative explanation, of course, is that mothers interpret these behaviors differently depending on whether they are exhibited by a son or a daughter (e.g., Condry & Condry, 1976). Mothers either downplay the severity/seriousness of these behaviors (i.e., attacks, fights) when they are manifested by boys, or they exaggerate it when manifested by girls, or both. Another explanation has to do with the conceptualization of physical aggression. The physically aggressive behaviors considered in this study are multidetermined; in other words, they likely involve a mix of various forms of physical aggression (e.g., reactive-defensive, proactive-offensive/hostile, instrumental). It could be that these behaviors (i.e., attacks, fights) represent a qualitatively different mix compared with the other three behaviors, with gender differences in the latter but not in the former. This is not a likely explanation, however, because we did not find evidence of "local dependence" among behavior items; that is, evidence of a direct relationship between behavior items beyond and above their separate relationship with the latent physical aggression variable.[4] In fact, the three-class model provided an excellent fit to the 17-month-old data on behavior items, suggesting that the behavior items are independent within each latent class. On the other hand, contrary to the gender paradox hypothesis, we did not find that high-aggressive 29-month-old girls were more likely than their male counterparts to have been so 1 year before.

Another Look at the Arrested Socialization Hypothesis

Some researchers have suggested that extreme antisocial behavior such as physical aggression is due to arrested socialization (e.g., Patterson, 1982; Tremblay, 2003). According to this view, aggressive children are simply children who have not grown up, have not learned not to be aggressive, with their behavior representing that which is normative for younger children. Contrary to arrested socialization hypothesis, our results suggest that there are more children drifting into than drifting out of manifesting physically aggressive behaviors on a frequent basis during toddlerhood. Moreover, a majority of 29-month-old aggressive girls had not manifested physically aggressive behaviors on a frequent basis 1 year earlier. Of course, this situation may be different in school-age children and adolescents, in which it is believed that there is a decreasing trend in the prevalence of physical aggression with age (e.g., Tremblay et al., 1996). Hence, it appears that this is indeed an oversimplified view of the development of aggressive behavior at least early in life, one that does not take into account the complex pattern of continuity and discontinuity in children's physical aggression status over time (Loeber & Stouthamer-Loeber, 1998).

On the other hand, our results suggest that physically aggressive toddlers display the same continuity in their (maladaptive) behaviors as their nonaggressive counterparts. One possible explanation is that the child's aggressive dispositions systematically select him or her into environments or life circumstances that further evoke and sustain physically aggressive behaviors against other children. In his paper on the effect of constitutional characteristics of children on parenting behavior, Bell (1968) proposed that highly assertive children (e.g., overtly aggressive) elicit less nurturance and more physical punishment from their parents because they can be controlled less adequately by love-oriented child-rearing techniques. In turn, these intruding and controlling efforts may generate frustration, conflict, and the establishment of a coercive cycle of violent interactions (Patterson, 1982). Hence, it appears that the processes that sustain physically aggressive behavior across time, and possibly situations, are being established very early on in life.

Prevalence of Physical Aggression Among Toddlers

Results from observational studies suggest that physical force is not used in a majority of toddlers' conflicts, which are otherwise relatively infrequent and short in duration (for a review, see Shantz, 1987). On the one hand, it seems that many toddlers use physical force at least once during these conflicts. For instance, Hay and Ross (1982) reported that over the course of 4 consecutive days (15 min of observation per day), 12 out of 48 toddlers who were 21 months old used force in struggles over objects. Similarly, in another study by Lewis and his colleagues (Lewis, Young, Brooks, & Michalson, 1975), peer-directed aggressive behaviors such as hitting, scratching, and biting were manifested by 22% of 12- to 18-month-old toddlers during a 15-min play session. Similarly, Bronson (1981) reported that 36 out of 40 toddlers engaged in "physical assaults" involving mainly pushing and hitting in three small-group play sessions around 17–20 months of age. These results are consistent with our results according to which a relatively large number of toddlers in the general population manifest physically aggressive behaviors on an occasional basis. On the other hand, our results suggest that it is possible to distinguish these toddlers from the relatively few others who manifest these behaviors on a frequent basis. Indeed, there is some evidence,

[3] The population attributable fraction was calculated as the proportion of physically aggressive boys in the population, minus that proportion had they been no more aggressive than girls over the proportion of aggressive children in the population (i.e., [.025 – .005]/.03 = .67).

[4] This is not to say, however, that the relative contribution of the various forms of physical aggression is the same for any given physically aggressive behavior. In fact, results from observational studies suggest that toddlers may manifest higher rates of one form than another (see Hay, 2005). Rather, whatever the relative contribution of these different forms of physical aggression may be, it is the same for all the physically aggressive behaviors considered in this study.

albeit limited, from observational studies of older toddlers suggesting that a minority of toddlers may be responsible for the majority of forceful actions. In a study of British toddlers, Hay et al. (2000) reported that 37 out of 66 toddlers who are 18, 24, and 30 months old used force proactively at home against familiar peers, but only a minority (i.e., 15 out of 37) did so more than once. In another study, Jersild and Markey (1935) observed 54 children ages 2–4 years during free play. Whereas 1 child was responsible for 87 attacks on peers, another child did not manifest any. Further, other studies suggest that forceful acts, such as starting a quarrel, may be stable characteristics of toddlers (Hay & Ross, 1982; see also Tremblay et al., 2004).

Studying Gender Differences in Aggressive Behavior Using a Latent Class Analysis Framework

Current research strategies in developmental psychopathology often involve using a cutoff point to distinguish nonaggressive from aggressive children. The groups of children created by this procedure are not homogeneous, however. In fact, some nonaggressive children will be classified as aggressive (i.e., false positives); conversely, some aggressive children will be classified as nonaggressive (i.e., false negatives). Classification errors are known to yield biased (under or over) estimates of the true prevalence, incidence, and remission (Baillargeon et al., 2004; Goldberg, 1975; Yanagawa & Gladen, 1984), thus making the comparison of these estimates between boys and girls hazardous. The latent class analysis framework adopted in this study provided maximum likelihood—asymptotically unbiased and efficient—estimates of the prevalence, incidence, and remission of physical aggression and of gender differences therein.

Limitations

This study is not without limitations, however. First, children born to mothers residing in Northern Québec, in Cree and Inuit territory, and on native reserves were not part of the target population. Although these exclusions represented only 2.1% of all live births to mothers residing in Québec, they may have resulted in slightly lower prevalence estimates, as mental health problems, at least in adults, seem more frequent among natives than in the rest of the Québec population (Daveluy, Lavallée, Clarkson, & Robinson, 1994; Jetté, 1994, 1995). Another factor that may have contributed to lower prevalence estimates is that certain premature babies (i.e., $<$ 24 weeks of gestation) were excluded from the first wave sample. However, these exclusions represented less than 0.1% of registered births at the time of data collection. Second, relying on only one informant is less than optimal, even though there is evidence that mothers are reliable for preschool-age children (Carter et al., 2003; Earls, 1980). Of course, aggressive behavior and gender differences therein may at least, in part, be context/informant specific. Moreover, the continuity in children's latent physical aggression status between 17 and 29 months of age may at least, in part, reflect continuity in mothers' perception of their children's behavior or reporting biases (e.g., social desirability, negative/positive response sets). Hence, further research is needed with more than one informant, and possibly observational assessments of the child's behavior, to test whether gender differences in the prevalence of aggressive behavior may be present in one setting but not another. More research is also needed to examine the predictive accuracy of early physical aggression when the informant reporting on the child's behavior varies from one time point to another. Note that these two limitations are fairly typical of large population-based surveys. Furthermore, the issue of whether early gender differences in aggressive behavior are due to boys being more vulnerable to risk factors than girls or boys being more exposed than girls to risk factors needs to be addressed (for a different view, see Plomin & Foch, 1981). Data from the QLSCD should help answer these and other important questions in the future.

References

Achenbach, T. M., Edelbrock, C. S., & Howell, C. T. (1987). Empirically based assessment of the behavioral/emotional problems of 2- and 3-year-old children. *Journal of Abnormal Child Psychology, 15,* 629–650.

Achenbach, T. M., & McConaughy, S. H. (1997). *Empirically based assessment of child and adolescent psychopathology: Practical applications* (2nd ed.). Thousand Oaks, CA: Sage.

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms and profiles.* Burlington: University of Vermont, Research Center for Children, Youth, and Families.

Anderson, J. E. (Chairman). (1972). *The young child in the home: A survey of three thousand American families—Report of the Committee on the Infant and Preschool Child.* In D. J. Rothman & S. M. Rothman (Series Eds.), *Family in America*. New York: Arno Press & The New York Times. (Original work published 1936)

Archer, J., & Côté, S. (2005). Sex differences in aggressive behavior: A developmental and evolutionary perspective. In R. E. Tremblay, W. W. Hartup, & J. Archer (Eds.), *Developmental origins of aggression* (pp. 425–443). New York: Guilford Press.

Baillargeon, R. H. (2002, July). *Gender differences in physical aggression at 17 months of age: Results from the Longitudinal Study of Child Development in Québec*. Paper presented at the 15th World Meeting of the International Society for Research on Aggression, Montréal, Québec, Canada.

Baillargeon, R. H., Normand, C. L., Séguin, J. R., Zoccolillo, M., Japel, C., Pérusse, D., et al. (in press). The evolution of problem and social competence behaviors during toddlerhood: A prospective population-based cohort survey. *Infant Mental Health Journal*.

Baillargeon, R. H., Tremblay, R. E., & Willms, D. (1999). *The prevalence of physical aggression in Canadian children: A multi-group latent class analysis of data from the First Collection Cycle (1994–1995) of the National Longitudinal Survey of Children and Youth (NLSCY).* Ottawa, Ontario, Canada: Applied Research Branch, Human Resources Development Canada. Retrieved December 1, 2005, from http://www.sdc.gc.ca/en/cs/sp/sdc/pkrf/publications/research/1999–001255/page00.shtml

Baillargeon, R. H., Tremblay, R. E., Willms, D., Romano, E., Lee, K., & Wu, H. -X. (2004). Modeling intraindividual change over time in the absence of a "gold standard." *Psychology Science, 46, 7*–35.

Baillargeon, R. H., Tremblay, R. E., Zoccolillo, M., Boivin, M., Pérusse, D., Japel, C., & Wu, H.-X. (2002). Intra-individual change in behaviour from 17 to 29 months of age. In *L'Étude longitudinale du développement des enfants du Québec* [Quebec Longitudinal Study of Children Development] (ELDEQ 1998–2002, Vol. 2, No. 7, pp. 1–40). Québec City, Québec, Canada: Institut de la Statistique du Québec. Retrieved December 1, 2005, from http://www.stat.gouv.qc.ca/publications/sante/bebe_v2no7_pdf_an.htm

Baillargeon, R. H., & Zoccolillo, M. (2003, April). *The evolution of disruptive behavior in toddlerhood: Results from the Québec Longitudinal Study of Child Development.* Paper presented at the biennial meeting of the Society for Research in Child Development, Tampa, FL.

Bakeman, R., & Brownlee, J. R. (1982). Social rules governing object conflicts in toddlers and preschoolers. In K. H. Rubin & H. S. Ross (Eds.), *Peer relationships and social skills in childhood* (pp. 99–111). New York: Springer-Verlag.

Behar, L. B., & Stringfield, S. (1974). A behavior rating scale for the preschool child. *Developmental Psychology, 10,* 601–610.

Bell, R. Q. (1968). A reinterpretation of the direction of effects in studies of socialization. *Psychological Review, 75,* 81–95.

Blumberg, S. J., Halfon, N., & Olson, L. M. (2004). The national survey of early childhood health. *Pediatrics, 113,* 1899–1906.

Bone, M., & Meltzer, H. (1989). *The prevalence of disability among children: OPCS surveys of disability in Great Britain, Report 3.* London: Social Survey Division, Office of Population Censuses and Surveys.

Bridges, K. M. B. (1933). A study of social development in early infancy. *Child Development, 4,* 36–49.

Briggs-Gowan, M. J., Carter, A. S., Skuban, E. M., & Horwitz, S. M. (2001). Prevalence of social–emotional and behavioral problems in a community sample of 1- and 2-year-old children. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 811–819.

Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, B., Dodge, K., et al. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. *Developmental Psychology, 39,* 222–245.

Bronson, W. C. (1981). *Toddlers' behaviors with agemates: Issues of interaction, cognition, and affect.* Norwood, NJ: Ablex.

Brownlee, J. R., & Bakeman, R. (1981). Hitting in toddler–peer interaction. *Child Development, 52,* 1076–1079.

Bühler, C. (1931). The social behavior of the child. In C. Murchison (Ed.), *A handbook of child psychology* (pp. 392–431). Worcester, MA: Clark University Press.

Bühler, C. (1935). *From birth to maturity: An outline of the psychological development of the child.* London: Routledge & Kegan Paul.

Bussey, K., & Bandura, A. (1999). Social–cognitive theory of gender development and differentiation. *Psychological Review, 106,* 676–713.

Cairns, R. B. (1979). *Social development: The origins and plasticity of interchanges.* San Francisco: Freeman.

Cairns, R. B., Cairns, B. D., Neckerman, H. J., Ferguson, L. L., & Gariépy, J. -L. (1989). Growth and aggression: 1. Childhood to early adolescence. *Developmental Psychology, 25,* 320–330.

Carter, A. S., Briggs-Gowan, M. J., Jones, S. M., & Little, T. D. (2003). The Infant–Toddler Social and Emotional Assessment (ITSEA): Factor structure, reliability, and validity. *Journal of Abnormal Child Psychology, 31,* 495–514.

Caspi, A., Elder, G. H., & Bem, D. J. (1987). Moving against the world: Life-course patterns of explosive children. *Developmental Psychology, 23,* 308–313.

Clogg, C. C. (1979). Some latent structure models for the analysis of Likert-type data. *Social Science Research, 8,* 287–310.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Condry, J., & Condry, S. (1976). Sex differences: A study in the eye of the beholder. *Child Development, 47,* 812–819.

Cressie, N., & Read, T. R. C. (1984). Multinomial goodness-of-fit tests. *Journal of the Royal Statistical Society, Series B, 46,* 440–464.

Daveluy, C., Lavallée, C., Clarkson, M., & Robinson, E. (Eds.). (1994). *A health profile of the Cree: Report of the Santé Québec Health Survey of the James Bay Cree 1991.* Québec City, Québec, Canada: Ministère de la Santé et des Services Sociaux, Gouvernement du Québec.

Desrosiers, H. (2000). Family, child care and neighborhood characteristics. In *Longitudinal Study of Child Development in Québec* (ELDEQ 1998–2002, Vol. 1, No. 2). Québec City, Québec, Canada: Institut de la Statistique du Québec. Retrieved December 1, 2005, from http://www.stat.gouv.qc.ca/publications/sante/bebe_no2_an.htm

Dunn, J., & Munn, P. (1985). Becoming a family member: Family conflict and the development of social understanding in the second year. *Child Development, 56,* 480–492.

Earls, F. (1980). The prevalence of behavior problems in 3-year-old children: Comparison of the reports of fathers and mothers. *Journal of the American Academy of Child Psychiatry, 19,* 439–452.

Earls, F. (1987). Sex differences in psychiatric disorders: Origins and developmental influences. *Psychiatric Developments, 1,* 1–23.

Eckerman, C. O., Whatley, J. L., & Kutz, S. L. (1975). Growth of social play with peers during the second year of life. *Developmental Psychology, 11,* 42–49.

Eme, R. F. (1992). Selective female affliction in the developmental disorders of childhood: A literature review. *Journal of Clinical Child Psychology, 21,* 354–364.

Fagot, B. I. (1984). The consequents of problem behavior in toddler children. *Journal of Abnormal Child Psychology, 12,* 385–396.

Fagot, B. I., & Hagan, R. (1985). Aggression in toddlers: Responses to assertive acts of boys and girls. *Sex Roles, 12,* 341–351.

Fienberg, S. E. (1980). *The analysis of cross-classified categorical data* (2nd ed.). Cambridge, MA: MIT Press.

Fleiss, J. L., Levin, B., & Paik, M. C. (2003). *Statistical methods for rates and proportions* (3rd ed.). Hoboken, NJ: Wiley.

Fowler, P. C., & Park, R. M. (1979). Factor structure of the Preschool Behavior Questionnaire in a normal population. *Psychological Reports, 45,* 599–606.

Glidewell, J. C., Mensh, I. N., & Gildea, M. C. -L. (1957). Behavior symptoms in children and degree of sickness. *American Journal of Psychiatry, 114,* 47–53.

Goldberg, J. D. (1975). The effects of misclassification on the bias in the difference between two proportions and the relative odds in the fourfold table. *Journal of the American Statistical Association, 70,* 561–567.

Golding, J., Pembrey, M., Jones, R., & ALSPAC Study Team. (2001). ALSPAC—The Avon Longitudinal Study of Parents and Children: 1. Study methodology. *Paediatric and Perinatal Epidemiology, 15,* 74–87.

Goodenough, F. (1931). *Anger in young children*. Westport, CT: Greenwood Press.

Goodman, L. A. (1981). Criteria for determining whether certain categories in a cross-classification table should be combined, with special reference to occupational categories in an occupational mobility table. *American Journal of Sociology, 87,* 612–650.

Hagenaars, J. A. (1993). *Loglinear models with latent variables.* London: Sage.

Hagenaars, J. A., & McCutcheon, A. L. (2002). *Applied latent class analysis.* Cambridge, England: Cambridge University Press.

Hay, D. F. (2005). The beginnings of aggression in infancy. In R. E. Tremblay, W. W. Hartup, & J. Archer (Eds.), *Developmental origins of aggression* (pp. 107–132). New York: Guilford Press.

Hay, D. F., Castle, J., & Davies, L. (2000). Toddlers' use of force against familiar peers: A precursor of serious aggression? *Child Development, 71,* 457–467.

Hay, D. F., & Ross, H. S. (1982). The social nature of early conflict. *Child Development, 53,* 105–113.

Heinstein, M. (1969). *Behavior problems of young children in California*. Berkeley: State of California, Bureau of Maternal and Child Health, Department of Public Health.

Holmberg, M. C. (1980). The development of social interchange patterns from 12 to 42 months. *Child Development, 51,* 448–456.

Hornberger, R. C., Bowman, J. C., Greenblatt, H. N., & Corsa, L., Jr. (1960). *Health supervision of young children in California.* Berkeley: State of California, Department of Public Health.

Huesmann, L. R. (2002). Commentary: Understanding aggression requires integrating diverse perspectives. *International Society for the Study of Behavioural Development Newsletter, 2*(42), 15–16.

Hyde, J. S. (1984). How large are gender differences in aggression? A developmental meta-analysis. *Developmental Psychology, 20,* 722–736.

Institut de la Statistique du Québec. (2002). *Le Québec statistique, édition 2002* [Statistics of Québec, 2002 edition]. Québec City, Québec, Canada: Gouvernement du Québec.

Janson, H. (2003a). Influences on participation rate in a Norwegian child development screening questionnaire study. *Acta Paediatrica, 92,* 91–96.

Janson, H. (2003b, April). *Parent-completed developmental screening in a Norwegian population sample: Comparison with U.S. data and sex differences.* Poster presented at the biennial meeting of the Society for Research in Child Development, Tampa, FL.

Jenkins, S., Bax, M., & Hart, H. (1980). Behaviour problems in pre-school children. *Journal of Child Psychology and Psychiatry, 21,* 5–17.

Jersild, A. T., & Markey, F. V. (1935). Conflicts between preschool children. In L. H. Meek (Ed.), *Child development monographs* (No. 21). New York: Columbia University, Teachers College.

Jetté, M. (Ed.). (1994). *A health profile of the Inuit: Report of the Santé Québec health survey among the Inuit of Nunavik, 1992* (Vols. I, II). Québec City, Québec, Canada: Ministère de la Santé et des Services Sociaux, Gouvernement du Québec.

Jetté, M. (Ed.). (1995). *A health profile of the Inuit: Report of the Santé Québec health survey among the Inuit of Nunavik, 1992* (Vol. III). Québec City, Québec, Canada: Ministère de la Santé et des Services Sociaux, Gouvernement du Québec.

Jetté, M. (2002). Survey description and methodology: Part I. Logistics and longitudinal data collections. In *Québec Longitudinal Study of Child Development* (QLSCD 1998–2000, Vol. 2, No. 1). Québec City, Québec, Canada: Institut de la Statistique du Québec. Retrieved December 1, 2005, from http://www.stat.gouv.qc.ca/publications/sante/bebe_v2no1_an.htm

Jetté, M., & Des Groseilliers, L. (2000). Survey description and methodology. In *Longitudinal study of child development in Québec* (ELDEQ 1998–2000, Vol. 1, No. 1). Québec City, Québec, Canada: Institut de la Statistique du Québec. Retrieved December 1, 2005, from http://www.stat.gouv.qc.ca/publications/sante/bebe_no1_pdf_an.htm

Kasmini, K., Kyaw, O., Krishnaswamy, S., Ramli, H., & Hassan, S. (1993). A prevalence survey of mental disorders among children in a rural Malaysian village. *Acta Psychiatrica Scandinavica, 87,* 253–257.

Keenan, K., Loeber, R., & Green, S. (1999). Conduct disorder in girls: A review of the literature. *Clinical Child and Family Psychology Review, 2,* 3–19.

Keenan, K., & Shaw, D. (1994). The development of aggression in toddlers: A study of low income families. *Journal of Abnormal Child Psychology, 22,* 53–77.

Keenan, K., & Shaw, D. (1997). Developmental and social influences on young girls' early problem behavior. *Psychological Bulletin, 121,* 95–113.

Knight, G. P., Fabes, R. A., & Higgins, D. A. (1996). Concerns about drawing causal inferences from meta-analyses: An example in the study of gender differences in aggression. *Psychological Bulletin, 119,* 410–421.

Lahey, B. B., Miller, T. L., Gordon, R. A., & Riley, A. W. (1999). Developmental epidemiology of the disruptive behavior disorders. In H. C. Quay & A. E. Hogan (Eds.), *Handbook of disruptive behavior disorders* (pp. 21–48). New York: Plenum.

Langeheine, R., & van de Pol, F. (1994). Discrete-time mixed Markov latent class models. In A. Dale & R. B. Davies (Eds.), *Analyzing social and political change: A casebook of methods* (pp. 170–197). London: Sage.

Lewis, M., Young, G., Brooks, J., & Michalson, L. (1975). The beginning of friendship. In M. Lewis & L. A. Rosenblum (Eds.), *Friendship and peer relations* (pp. 27–65). New York: Wiley.

Links, P. S. (1983). Community surveys of the prevalence of childhood psychiatric disorders: A review. *Child Development, 54,* 531–548.

Loeber, R., & Hay, D. F. (1997). Key issues in the development of aggression and violence from childhood to early adulthood. *Annual Review of Psychology, 48,* 371–410.

Loeber, R., & Stouthamer-Loeber, M. (1998). Development of juvenile aggression and violence: Some common misconceptions and controversies. *American Psychologist, 53,* 242–259.

Lytton, H. (1990). Child and parent effects in boys' conduct disorder: A reinterpretation. *Developmental Psychology, 26,* 683–697.

Lytton, H., & Romney, D. M. (1991). Parents' differential socialization of boys and girls: A meta-analysis. *Psychological Bulletin, 109,* 267–296.

Maccoby, E. E. (2002). Perspectives on gender development. In W. W. Hartup & R. K. Silbereisen (Eds.), *Growing points in developmental science: An introduction* (pp. 202–222). New York: Psychology Press.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

Maccoby, E. E., & Jacklin, C. N. (1980). Sex differences in aggression: A rejoinder. *Child Development, 51,* 964–980.

Macfarlane, J. W., Allen, L., & Honzik, M. P. (1954). *A developmental study of behavior problems in normal children*. Berkeley: University of California Press.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47,* 149–174.

Mathiesen, K. S., & Sanson, A. (2000). Dimensions of early childhood behavior problems: Stability and predictors of change from 18 to 30 months. *Journal of Abnormal Child Psychology, 28,* 15–31.

Maudry, M., & Nekula, M. (1939). Social relations between children of the same age during the first two years of life. *Journal of Genetic Psychology, 54,* 193–215.

McCutcheon, A. L. (1987). *Latent class analysis.* London: Sage.

Moffitt, T. E., Caspi, A., Rutter, M., & Silva, P. A. (2001). *Sex differences in antisocial behaviour: Conduct disorder, delinquency, and violence in the Dunedin longitudinal study*. Cambridge, England: Cambridge University Press.

Patterson, G. R. (1982). *A social learning approach: Vol. 3. Coercive family process*. Eugene, OR: Castalia.

Plante, N., Courtemanche, R., & Des Groseilliers, L. (2002). Survey description and methodology: Part II. Methodology—Longitudinal aspects of the first three rounds 1998 to 2000). In *Québec Longitudinal Study of Child Development* (QLSCD 1998–2000, Vol. 2, No. 1). Québec City, Québec, Canada: Institut de la Statistique du Québec. Retrieved December 1, 2005, from http://www.stat.gouv.qc.ca/publications/sante/bebe_v2no1_an.htm

Plomin, R., & Foch, T. T. (1981). Sex differences and individual differences. *Child Development, 52,* 383–385.

Raine, A. (2002a). Annotation: The role of prefrontal deficits, low autonomic arousal, and early health factors in the development of antisocial and aggressive behavior in children. *Journal of Child Psychology and Psychiatry, 43,* 417–434.

Raine, A. (2002b). Biosocial studies of antisocial and violent behavior in children and adults: A review. *Journal of Abnormal Psychology, 30,* 311–326.

Restoin, A., Montagner, H., Rodriguez, D., Girardot, J. J., Laurent, D., Kontar, F., et al. (1985). Chronologie des comportements de communication et profils de comportement chez le jeune enfant [Chronology of communication behaviors and profiles of behavior in young children]. In R. E. Tremblay, M. A. Provost, & F. F. Strayer (Eds.), *Ethologie et développement de l'enfant* [Ethology and child development] (pp. 93–130). Paris: Stock/Laurence Pernoud.

Rossi, J. S. (1983). Ratios exaggerate gender differences in mathematical ability. *American Psychologist, 38,* 348.

Rost, J., & Langeheine, R. (1997). *Applications of latent trait and latent class models in the social sciences.* Berlin: Waxman.

Rutter, M. (1967). A children's behaviour questionnaire for completion by teachers: Preliminary findings. *Journal of Child Psychology and Psychiatry, 8,* 1–11.

Shantz, C. U. (1987). Conflicts between children. *Child Development, 58,* 283–305.

Shaw, D., Keenan, K., & Vondra, J. I. (1994). Developmental precursors of externalizing behavior: Ages 1 to 3. *Developmental Psychology, 30,* 355–364.

Shirley, M. M. (1933). *The first two years—A study of twenty-five babies: Vol. II. Intellectual development*. Westport, CT: Greenwood Press.

Silverthorn, P., & Frick, P. J. (1999). Developmental pathways to antisocial behaviour: The delayed-onset pathway in girls. *Development and Psychopathology, 11,* 101–126.

Smetana, J. G. (1989). Toddlers' social interactions in the context of moral and conventional transgressions in the home. *Developmental Psychology, 25,* 499–508.

Stevenson, J., & Goodman, R. (2001). Association between behaviour at age 3 years and adult criminality. *British Journal of Psychiatry, 179,* 197–202.

Thissen, D., Steinberg, L., & Gerrard, M. (1986). Beyond group-mean differences: The concept of item bias. *Psychological Bulletin, 99,* 118–128.

Tieger, T. (1980). On the biological basis of sex differences in aggression. *Child Development, 51,* 943–963.

Tremblay, R. E. (2003). Why socialization fails: The case of chronic physical aggression. In B. B. Lahey, T. E. Moffitt, & A. Caspi (Eds.), *Causes of conduct disorder and juvenile delinquency* (pp. 182–224). New York: Guilford Press.

Tremblay, R. E., Boulerice, B., Harden, P. W., McDuff, P., Pérusse, D., Pihl, R. O., & Zoccolillo, M. (1996). Do children in Canada become more aggressive as they approach adolescence? In Human Resources Development Canada & Statistics Canada (Eds.), *Growing up in Canada: National Longitudinal Survey of Children and Youth* (pp. 127–137). Ottawa, Ontario, Canada: Statistics Canada.

Tremblay, R. E., Japel, C., Pérusse, D., McDuff, P., Boivin, M., Zoccolillo, M., & Montplaisir, J. (1999). The search for the age of 'onset' of physical aggression: Rousseau and Bandura revisited. *Criminal Behaviour and Mental Health, 9,* 8–23.

Tremblay, R. E., Nagin, D. S., Séguin, J. R., Zoccolillo, M., Zelazo, P. D., Boivin, M., et al. (2004). Physical aggression during early childhood: Trajectories and predictors. *Pediatrics, 114*(1), 43–50.

U.S. National Health Survey. (1959). *Children and youth. Selected health characteristics, United States, July 1957–June 1958.* Washington, DC: U.S. Department of Health, Education, and Welfare, Public Health Service, Division of Public Health Methods.

van de Pol, F., & Langeheine, R. (1990). Mixed Markov latent class models. In C. C. Clogg (Ed.), *Sociological methodology 1990* (pp. 213–247). Oxford, England: Blackwell.

Vermunt, J. K. (1997, September 18). lEM: A general program for the analysis of categorical data [Computer program]. Tilburg, the Netherlands: Tilburg University.

von Eye, A., & Clogg, C. C. (1994). *Latent variables analysis: Application for developmental research*. London: Sage.

Yanagawa, T., & Gladen, B. C. (1984). Estimating disease rates from a diagnostic test. *American Journal of Epidemiology, 119,* 1015–1023.

Zahn-Waxler, C. (1993). Warriors and worriers: Gender and psychopathology. *Development and Psychopathology, 5,* 79–89.

Zoccolillo, M. (1993). Gender and the development of conduct disorder. *Development and Psychopathology, 5,* 65–78.

Received August 3, 2004
Revision received November 9, 2005
Accepted November 29, 2005 ■