Journal of Abnormal Psychology
2004, Vol. 113, No. 2, 258–271

0021-843X/04/$12.00 DOI: 10.1037/0021-843X.113.2.258

Clinically Abusive Relationships in an Unselected Birth Cohort: Men's and Women's Participation and Developmental Antecedents

Miriam K. Ehrensaft
Columbia University and New York State Psychiatric Institute

Terrie E. Moffitt and Avshalom Caspi
King's College and University of Wisconsin—Madison

In an unselected birth cohort (N = 980, age 24–26 years), individuals in abusive relationships causing injury and/or official intervention (9% prevalence) were compared with participants reporting physical abuse without clinical consequences and with control participants who reported no abuse, on current characteristics and prospective developmental risks. In nonclinically abusive relationships, perpetrators were primarily women. In clinically abusive relationships, men and women used physical abuse, although more women needed medical treatment for injury. Women in clinically abusive relationships had childhood family adversity, adolescent conduct problems, and aggressive personality; men had disinhibitory psychopathology since childhood and extensive personality deviance. These findings counter the assumption that if clinical abuse was ascertained in epidemiological samples, it would be primarily man-to-woman, explained by patriarchy rather than psychopathology.

Differing theoretical (and political) perspectives about partner abuse have generated heated debates about the definition and etiology of partner abuse. Principal among these is whether studies of community samples, particularly those that are based in survey methodology, can detect cases of so-called "real" abuse, similar to that observed in samples of women seeking help at a battered women's shelter, police precinct, or local emergency room. A second debate focuses on the direction (e.g., man to woman) and degree of mutuality of abuse, and a third pertains to whether men and women's partner abuse is rooted in psychopathology. This study uses a longitudinal research design with an unselected birth cohort to test hypotheses regarding these controversies.

Miriam K. Ehrensaft, Division of Child Psychiatry, Columbia University College of Physicians and Surgeons and New York State Psychiatric Institute; Terrie E. Moffitt and Avshalom Caspi, Institute of Psychiatry, King's College London, London, United Kingdom, and Department of Psychology, University of Wisconsin—Madison.

We thank the Dunedin Study members, their parents, and teachers; Dunedin Unit Director Richie Poulton; Unit research staff; and Study founder Phil Silva. For official crime data, we thank Paul Stevenson and the Dunedin Police. Honalee Harrington and Helena Kiernan provided research assistance. The clients and staff of Arc Community Corrections of Madison, Wisconsin, supported our reliability studies of the Partner Conflict Calendar. The Dunedin Multidisciplinary Health and Development Research Unit is supported by the New Zealand Health Research Council. This research received support from Grant K08-MH01913-01 from the U.S. National Institute of Mental Health to Miriam Ehrensaft and from U.S. National Institute of Mental Health Grants MH45070, MH49414, and MH56344; the Medical Research Council; a Royal Society-Wolfson Merit Award; and Air New Zealand.

Correspondence concerning this article should be addressed to Miriam K. Ehrensaft, Division of Child Psychiatry, Columbia University, College of Physicians and Surgeons, 1051 Riverside Drive, Unit 78, New York, NY 10032, or to Terrie E. Moffitt, Box Number P080, SGDP Centre, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, United Kingdom. E-mail: ehrensam@childpsych.columbia.edu or t.moffitt@iop.kcl.ac.uk

Can "Real" Abuse Be Detected in Community Samples?

Some suggest that divergent theories of partner abuse are the product of studying qualitatively different samples of abusers and victims. The most widely cited advocate of this feminist-based perspective is Johnson (1995), who argued that researchers ought to consider participants derived from community surveys versus battered women's shelters as representing "largely nonoverlapping populations, experiencing different forms of violence" (p. 289). According to this concept, survey populations only detect cases of *common couple violence*, which does not result in serious injury or come to the attention of emergency personnel or other clinicians. Studies that are based on samples of women from shelters, Johnson argued, are representative of significant violence and this man-to-woman violence is driven by *patriarchal terrorism*, culturally sanctioned violence intended to force females' compliance. Later, Johnson updated his terminology to accommodate accumulating evidence of concurrent psychopathology in male batterers (for review see Holtzworth-Munroe, Meehan, Herron, Rehman & Stuart, 2000); *situational violence* denotes the less serious forms of abuse, and *intimate terrorism* denotes more systematically abusive behaviors motivated by the intent to control a partner (Johnson & Ferraro, 2000). Straus (1999), who developed the most widely used survey instrument of family violence, known as the Conflict Tactic Scales (CTS; Straus, 1990), supported this position. He argued that, because of profound methodological differences, findings from samples of women in shelter and emergency room settings should not be generalized to cases of assault reported in family conflict surveys. Straus (1999) also proposed that the former methodology is essential for informing the prevention of more serious types of domestic abuse, whereas the latter should inform primary prevention of '"minor" partner violence.

The argument that community samples cannot represent clinically significant forms of domestic violence is problematic for several reasons. First, the argument contradicts epidemiological methodology by stating that real cases of a clinically significant

phenomenon cannot be found with well-designed epidemiological research. On the contrary, epidemiological methods reliably detect many important clinical phenomena, including HIV (Buve et al., 2001) and substance dependence (Frischer, Hickman, Kraus, Mariani, & Wiessin, 2001), which share with the study of partner violence complex ethical and methodological obstacles to case ascertainment. Well-designed longitudinal studies that enroll unselected cohorts known to represent populations and follow them without attrition bias should represent people involved in clinically significant domestic violence. Second, many victims seriously injured and psychologically intimidated by a current or former intimate partner avoid police, social service involvement, and shelters (Gondolf, 1998). Even when victims do seek refuge in shelters, a high degree of selectivity often influences which applicants are admitted, such as exclusion of aggressive women (Loseke, 1992). Thus, samples recruited from shelters and police records are not representative of serious partner violence, which raises questions about limiting the definition of *clinically important* partner violence to those so identified.

Epidemiological principles would suggest that unselected cohort samples contain shelter, emergency room, and forensic domestic violence cases as a fraction within the cohort that represents their prevalence in the general population, plus clinically significant forms of serious domestic violence that elude official attention, thus representing clinically significant domestic violence without selection bias. In the special case where a cohort has been followed longitudinally and assessed repeatedly, participants are able to provide unusually valid self-reports, because they develop confidence in the confidentiality guarantee that exceeds the confidence of naïve research participants or participants involved in legal proceedings related to domestic violence. The present study compared clinically versus nonclinically defined cases, within a single cohort sample, to test the hypothesis that they differ on perpetrators' sex and developmental antecedents. The following sections discuss these two controversies.

Is Real Abuse Mutual?

Findings of woman-to-man domestic violence from community surveys (Archer, 2000) have been criticized by practitioners and researchers as describing only trivial, common couples abuse, with the assumption that clinically defined abuse would be mainly man-to-woman violence (Johnson & Ferraro, 2000; O'Leary, 2000). Feminist theories propose that violence in intimate, heterosexual relationships is distinct from other forms of aggression and antisocial behavior because it occurs in the context of gender-based, institutionalized power asymmetry (Dobash & Dobash, 1992; Koss et al., 1994). This perspective cites data from crime and personal safety surveys, which uniformly find elevated ratios of male to female offenders of intimate partner violence and higher rates of injuries to women than men (Bachman & Saltzman, 1995; Tjaden & Thoennes, 1998; Zlotnick, Kohn, Peterson, & Pearlstein, 1998).

Feminist theory, however, has been criticized as incomplete, because it cannot explain individual differences in the degree and frequency of partner abuse (Magdol, Moffitt, Caspi, & Silva, 1998). Recent studies do not confirm this assumption that real abuse is only man to woman. In prior work with the Dunedin Multidisciplinary Health and Development Study, a couples analysis of partner abuse found evidence for a dyadic process model of partner abuse, in which both partners' negative emotionality independently and jointly predicted couples' reciprocal abuse, even for severe cases (Moffitt, Robins, & Caspi, 2001). Capaldi and Owen's (2001) study of a community sample found that frequent physical aggression toward a partner was bidirectional, rather than being male only or female only. Types and frequencies of abusive acts tend not to differ significantly by sex, although their physical consequences are far greater for females (Straus, 1999).

Moreover, important methodological differences in partner abuse studies using crime victimization surveys versus family conflict surveys call into question our current knowledge about the directionality of clinical and nonclinical abuse. This point was cogently argued by Straus (1999). Briefly, crime surveys find much higher rates of male-to-female assault, whereas family conflict survey methodology reveals equal rates of assault by the two sexes. Crime surveys ask respondents to recall threats to personal safety, which tend to be more rare and extreme, whereas family conflict surveys ask participants to recall a broader range of behaviors used during the context of dyadic conflict. Moreover, crime surveys include assaults by former partners, whereas most family conflict studies of partner violence do not. Because the rate of assaults by former partners is higher than assaults by current partners (Bachman & Saltzman, 1995), Straus argued that these differing methodological approaches yielded qualitatively distinct conclusions about directionality. Last, even in crime victimization surveys, the high rates of male-to-female assault emerge from lifetime recall reports, whereas past-year reports, which are known to be more reliable, reveal more female-to-male assault (Tjaden & Thoennes, 1998). In summary, our interpretation of the existing literature is that the jury is still out on whether partner abuse, both clinical and nonclinical, is primarily man to woman versus bidirectional.

Is Real Abuse Psychopathological?

A third debate concerns whether partner abuse is rooted in psychopathology. Although feminist theorists allow that some men who chronically terrorize their partners with severe, controlling abuse may have a broader pattern of antisocial psychopathology (Johnson & Ferraro, 2000), the theory considers psychopathology as secondary. When feminist theorists do address psychopathology, it is almost always described as a woman's traumatic reaction to a male's abuse (e.g., Walker, 1989).

On the other hand, developmental theorists would predict that real abuse is pathological, because they consider partner violence to be a developmental expression of a continuous pattern of antisocial behavior over the life span (Farrington & West, 1990; Huesmann, Eron, Lefkowitz, & Walder, 1984). These theorists propose that aggressive, hostile styles of behavior, emotion regulation, cognition, and conflict resolution are first shaped by parent–child and interparental relationships and are then generalized across future development periods to the peer network and, later, to the individual's romantic relationships (Capaldi & Clark, 1998; Ehrensaft, Cohen, et al., 2003; O'Leary, 1999).

Evidence suggests that the perpetrators of partner violence having clinical consequences should have higher levels of preexisting psychopathology. First, adult men who perpetrate severe, chronic

partner abuse are characterized by concurrent, long-standing antisocial behavior and emotion dysregulation (Holtzworth-Munroe et al., 2000; Moffitt, Robins, et al., 2001). Consistent with the model of partner violence as a developmental expression of antisocial behavior, several large, epidemiological studies concur that prospectively measured early antisocial behavior strongly predicts partner violence in early adulthood (Capaldi & Clark, 1998; Ehrensaft, Cohen, et al., 2003; Magdol et al., 1998). Furthermore, prospectively measured risks for partner abuse, including socioeconomic resources, poor family relations, educational achievement, and problem behavior, are quite similar to those for other forms of antisocial behavior (e.g., Magdol et al., 1998). Like general physical aggression and antisocial behavior, partner violence is most prevalent in the period of adolescence to early adulthood (O'Leary, 1999).

The feminist view further maintains that women's partner abuse is not driven by psychopathology. However, accumulating evidence from prospective studies suggests that women's partner abuse may be associated with antisocial behavior in childhood and adolescence (Andrews, Foster, Capaldi, & Hops, 2000; Ehrensaft, Wasserman, et al., 2003; Giordano, Millhollin, Cernkovich, Pugh, & Rudolph, 1999; Magdol et al., 1998). Furthermore, there is evidence of assortative, or selective, partnering of men with antisocial histories and women with depressive and antisocial histories, which in turn increases the risk for partner abuse by both sexes (Kim & Capaldi, in press; Moffitt, Caspi, Rutter, & Silva, 2001). Last, both men and women who engaged in frequent and bidirectional physical aggression had higher levels of antisocial behavior, compared with men and women who perpetrated lower levels of partner aggression (Capaldi & Owen, 2001). Such findings oppose the traditional view that a preexisting history of psychopathology is not relevant to partner abuse etiology among women in highly abusive relationships. To date, sex comparisons in common versus clinical abuse have not been reported.

Personality development offers a useful conceptual framework for studying the link between partner abuse and pathology. This developmental perspective on individual differences suggests that problem behaviors, including general crime and partner violence, are manifestations of personality (Krueger, Caspi, & Moffitt, 2000; Krueger, Schmutte, Caspi, & Moffitt, 1994). In fact, research is converging to suggest that personality dysfunction may characterize some of the most serious male abusers (e.g., for review, see Holtzworth-Munroe et al., 2000; Holtzworth-Munroe & Stuart, 1994).

What is missing currently is a research base linking personality development and clinically defined partner violence, based on prospective research designs with representative samples of men and women, that tests whether personality deviance predates partner violence. Previous research with the present sample sheds some light on this issue. High levels of a negative emotionality personality trait were found to be shared by both generally antisocial and partner-violent men and women (Moffitt, Krueger, Caspi, & Fagan, 2000). The present study attempted to address the heterogeneity of partner abuse. To do so, we compared men and women involved in clinically versus nonclinically defined abuse on a broader set of measures of psychopathology and personality, taken many years earlier.

Study Goals and Hypotheses

In this study we sought to empirically address the three key theoretical debates identified above. First, we tested whether clinically defined partner abuse can be detected in an epidemiological sample of young men and women. Second, we tested the field's hypothesis that clinically defined abuse is man-to-woman, not woman-to-man. We hypothesized that this would not be the case and instead expected that clinically defined abuse would involve abuse by both men and women. Third, we aimed to test whether a prior history of psychopathology is a risk for men but not women in clinically defined abusive relationships. We expected that similar risk patterns of early child-rearing, family relations, undercontrolled behavior, and deviant personality functioning would exist for men and women in clinically abusive relationships, on the basis of the prior studies reviewed above.

Method

Participants and Procedure

Participants were members of the Dunedin Multidisciplinary Health and Development Study, an ongoing longitudinal investigation of health and behavior (Moffitt, Caspi, et al., 2001; Silva & Stanton, 1996). The cohort of 1,037 children (52% male, 48% female) was constituted at age 3 when the investigators enrolled 91% of the consecutive births between April 1972 and March 1973 in Dunedin, New Zealand. Cohort families represent the full range of socioeconomic status in the general population of New Zealand's south island and they are primarily White; fewer than 9% self-identified at age 26 as Maori or Pacific Islanders. The Dunedin sample has been assessed with a broad battery of psychological, medical, and sociological measures with high rates of participation. This report uses data from assessments of individuals at the following time points: 3 years (n = 1,037), 5 years (n = 991), 7 years (n = 954), 9 years (n = 955), 11 years (n = 925), 13 years (n = 850), 15 years (n = 976), 18 years (n = 993), and 26 years (n = 980; 96% of the living cohort members). (Prior partner abuse reports from the Dunedin Study have used abuse data collected at age 21). Those not followed up have not differed on antisocial behaviors from those followed up (Moffitt, Caspi, et al., 2001).

The research procedure involved bringing 4 study members per day (including emigrants living overseas) to the research unit within 60 days of their birthday for a full day of individual data collection. Each research topic was presented, in private, as a standardized module by a different trained examiner in counterbalanced order throughout the day. Data were also gathered from courts, parents, teachers, and informants who knew the study members well. Interviewers were blind to all data about the study members, and had tertiary degrees and professional experience in social work, medicine, public health, or clinical psychology.

Measures

Assessment of Developmental Risk Factors for Partner Violence

On the basis of prior research and theory, we selected five types of candidate risk on which to compare groups who were in clinically abusive, nonclinically abusive, and nonabusive relationships: family-of-origin characteristics, parenting, child behavior problems, adolescent psychiatric disorders, and adolescent personality traits. All of the risk-factor measures in this article have published evidence of their reliability and validity in the Dunedin Study, and these psychometric details are summarized in a book (Moffitt, Caspi, et al., 2001).

Assessment of Family-of-Origin Characteristics

Socioeconomic status (SES). SES was measured on a 6-point scale ranging from 1 (*unskilled*) to 6 (*professional*) that categorized each occupation into one of six groups on the basis of the education and income level associated with that occupation in the New Zealand census data (Elley & Irving, 1976). Measures of parental SES from ages 1 to 15 years were averaged ($\alpha = .92$). Lower scores imply more risk.

Mother's age at her first birth. This scale measured the age of the study member's mother when her first child was born, regardless of her age when the study child was born. Lower scores imply more risk.

Number of caretaker changes experienced by the child. This was assessed by summing the number of parent-figure changes (range: 0–6) each study member experienced from birth to age 11. At each assessment year, the parents were asked about changes in the family configuration since the last assessment. Responses included parent death, separation, cohabitation, remarriage, child sent to relatives, or foster care. Higher scores imply more risk.

Years with a single parent. This scale measures the number of years from birth to age 11 that the study member lived with a single parent. Higher scores imply more risk.

Assessment of Parenting

Negative mother–child interaction. At age 3, during a 1-hr testing session and during the child's physical examination, the mother's treatment of her child was observed and coded on eight categories by psychometricians and physicians. The scale designates 1 point for each behavior rated as rejecting or neglectful (e.g., mother's affect toward the child was consistently negative, or mother was unresponsive to child's needs, $\alpha = .71$). Higher scores imply more risk.

Harsh discipline. Harsh discipline was measured at ages 7 and 9 using a checklist of disciplinary behaviors on which mothers indicated if they engaged in 10 behaviors, such as "smack [your child] or hit him/her with something," "try to frighten [your child] with someone like his/her father or a policeman," and "threaten to smack, or deprive [your child] of something." Higher scores imply more risk.

Inconsistent discipline. This scale was administered at ages 7 and 9 as part of an interview about ways mothers coped with the study child when he or she misbehaved. Mothers evaluated their own and their husband's discipline on a 4-point scale (1 = *always the same*; 4 = *very changeable*) and the responses were summed. Higher scores imply more risk.

Assessment of Child Behavior Problems

At ages 5, 7, 9, and 11 years, parents and teachers of study members completed the Rutter Child Behavior Scale (Elander & Rutter, 1996). In the present study we examined the Hyperactive and Antisocial Behavior problem scales. Items were scored on a 3-point scale (0 = *doesn't apply*; 2 = *certainly applies*). Scores on these scales were averaged over the four time points (all αs $> .70$), to produce separate scales for parents and teachers. Higher scores imply more pathology.

Assessment of Adolescent Psychiatric Disorders

At ages 11, 13, and 15 years, study members were administered the Diagnostic Interview Schedule for Children (DISC-C; Costello, Edelbrock, Kalas, Kessler, & Klaric, 1982). Psychiatric disorders were diagnosed according to *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.; *DSM-III*; American Psychiatric Association, 1987) criteria. The DISC-C diagnoses have shown good interrater reliability in this cohort ($\kappa > .86$).

Assessment of Adolescent Personality Profiles

At age 18 years, the Multidimensional Personality Questionnaire (MPQ; Patrick, Curtin, & Tellegen, 2002), one of the best-known contemporary structural models of personality (Church & Burke, 1994), was administered to study members. This instrument was developed and standardized with nonclinical populations, has established reliability and validity, and yields a comprehensive profile of personality dimensions that are partially heritable, predictable from childhood, and very stable from adolescence to adulthood (for psychometric details, see Moffitt, Robins, et al., 2001). The MPQ provides scores on 10 distinct personality traits. High scores on the Self-Control scale (20 items, $\alpha = .79$) describe someone who is reflective, cautious, careful, rational, and plans ahead. High scores on the Harm Avoidance scale (21 items, $\alpha = .71$) describe someone who avoids excitement and danger and prefers safe activities even if they are tedious. High scores on the Traditionalism scale (22 items, $\alpha = .63$) describe individuals who desire a conservative social environment and endorse high moral standards. High scores on the Stress Reaction scale (14 items, $\alpha = .80$) describe someone who is nervous, vulnerable, sensitive, and prone to worry. High scorers on the Alienation scale (17 items, $\alpha = .76$) expect mistreatment and betrayal and are suspicious of others. High scorers on the Aggression scale (18 items, $\alpha = .78$) are willing to hurt others for their own advantage and frighten and cause discomfort for others. High scorers on the Well-Being scale (11 items, $\alpha = .67$) have a happy, cheerful disposition, feel good about themselves, and see a bright future. High scorers on the Social Potency scale (18 items, $\alpha = .76$) are forceful and decisive, fond of influencing others, and enjoy leadership roles. High scorers on Achievement (17 items, $\alpha = .69$) work hard and enjoy demanding projects and working long hours. Last, those who score high on Social Closeness (19 items, $\alpha = .75$) are sociable, affectionate, like people, and turn to others for comfort.

Measures of Partner Violence

This research used two distinct measures of partner abuse. The Dunedin Study Abuse scales (similar to the CTS) were used to ascertain the variety of different abuse acts performed by the respondent and his or her partner in the respondent's current or most recent relationship. The Partner Conflict Calendar (PCC; Caspi et al., 1996) was used to ascertain the consequences of abuse in respondents' relationships that were abusive. The relationships described in the Dunedin Study Abuse scales lasted on average 3.3 years prior to the interview ($SD = 2.7$), whereas the PCC covered the 3 years prior to the interview. As such, the reporting periods overlapped for almost all cases, that is, the scales and calendar assessed the same relationship.

The PCC. Study members (ages 24–26 years inclusive) reported their histories of abuse during the past 3 years, using the PCC. Early in the interview day, each study member had completed the Life History Calendar (LHC; see Caspi et al., 1996), a visual data-collection grid chart for obtaining retrospective reports of life events. The rows of the LHC grid refer to activity lines (e.g., residential histories, cohabitation histories, marriages, separations, births, child rearing, family deaths, work and unemployment trajectories, and education experiences). The columns divide the grid into years and months. The interviewer and respondent work together to fill in the cells of the grid. The "calendar" method capitalizes on advances in survey methodology and the cognitive psychology of memory to collect reliable retrospective data: It is a visual aid, it inquires about streams of events rather than isolated events, and it contextualizes questions about life events by linking them to other events. Reliability of recall is tested by comparing LHC retrospective reports (e.g. age-21 events reported at age 26) versus the gold-standard contemporaneous reports (age-21 events reported at age 21), and we and others have found excellent reliability, with above 90% agreement (Belli, Shay, & Stafford, 2001; Caspi et al., 1996).

Later in the same assessment day, the PCC was laid alongside the study member's completed LHC to cue memories about any months during the past 3 years in which violent incidents had occurred. A violent incident was defined by showing the respondent a card with "pushing, shoving, twisting

arm, grabbing, shaking, throwing, choking, strangling, using a knife or gun, kicking, biting, punching and slapping" in large block letters. Respondents were asked to recall months when any of these things happened "between you and the partner" (the PCC did not ascertain whether the male or female partner perpetrated violent acts, as we ascertained this using the physical abuse scales, see below). Interviewers facilitated recall of the timing of abuse events by asking questions derived from the study member's LHC such as, "Was that when you were working at ___?"; "Was that before or after your baby was born?"; or "Was that when you were living with ___, or after you moved out?" After all months with violence were recorded, interviewers asked follow-up questions about each month to ascertain how many incidents occurred during the month, whether either partner was injured in that month, whether treatment was sought, whether either partner was intoxicated by drugs or alcohol at the time of the violent incident, and whether agencies became involved (police, shelter or refuge, counselor or therapist, lawyers or courts).

We had earlier evaluated the reliability of the PCC by administering it on two occasions, 1 month apart, to 24 female inmates in the Wisconsin women's community corrections system. Those women reported abuse in an average of 8 months of their 36-month calendars, totaling an average of 60 violent incidents per woman. Although the female inmates had extensive substance abuse histories, low cognitive ability, and complex histories of domestic violence, their 1-month test–retest agreement was very good when aided by the calendar method, yielding kappas from .6 to 1.0 for the precise timing of months with violence. The test–retest correlation for "total number of incidents per month" was .43, but test–retest correlations were above .80 for number of months with violence, number of months with injury, number of months with medical treatment, number of months when substances were used prior to abuse, and number of months with agency involvement. The PCC, developed with National Institute of Mental Health support, has not been reported previously.

The Dunedin Study Abuse scales. These scales were used to assess specific abusive behaviors from one intimate partner to the other at age 26 (details in Magdol et al., 1998; Moffitt et al., 1997). Following Hornung, McCullough, and Sugimoto (1981), study members answered questions twice, first about their behavior toward their current or most recent partner and second about the partner's behavior toward them. The measure yields separate scores for perpetration and victimization, for both male and female study members. To avoid literacy problems, the interviewer read the questions aloud, and the respondent answered on a private answer sheet. The Physical Abuse scales contain all nine items of Straus's (1990) Conflict Tactics Scales (CTS, Form R; e.g., slap, choke, beat up), plus four items describing other physically abusive acts (e.g., twisting arm, bodily throw). A variety scale score was used here, summing the number of different types of abusive acts. Variety scales are extensively used to measure violence in the field of criminology and are highly correlated with frequency and seriousness measures, but they have stronger reliability and predictive validity than frequency or seriousness measures (Robins, 1978). All reliability αs were over .70 for men's and women's reports of perpetration and victimization. The content of the scale was described in this journal (Magdol et al., 1998).

Previous research with the scales in this cohort has demonstrated that both men's and women's reports explain more than three quarters of the variance in their partner's reports, showing that couple members corroborate each others' self-reports (Moffitt et al., 1997). Agreement was good because study members have learned, over 15 years of interview assessments about highly sensitive topics, to trust the confidentiality guarantee and, because they were not recruited to participate on the basis of their abuse involvement, they had little reason to bias their reports. Some research teams have found poor agreement between partners' reports of abuse, but these have tended to be from samples drawn in clinical or justice-system settings (where partners have a vested interest in how they portray their roles in abuse) and in samples interviewed for the first time (where respondents have no reason to trust the confidentiality guarantee). Such studies have recommended the use of both partners' reports, on the basis of the unreliability of partners' reports in such studies. In contrast, reliance on one partner's report is empirically supported for the Dunedin Study.

We included the Controlling Abuse Scale to assess contextual information about physical abuse; previous reports indicate that coercive control is a key motivator in cases of serious and chronic partner abuse (Johnson & Ferraro, 2000). The 10 items on this scale can be differentiated from the more general umbrella structure of psychological abuse, which includes both tactics having the intent or effect of controlling the partner and tactics that are unpleasant and abusive but not controlling (Ehrensaft, Langhinrichsen-Rohling, Heyman, O'Leary, & Lawrence, 1999; Ehrensaft & Vivian, 1999). For example, telling a partner that he or she could not work or study, stopping a partner from seeing family or friends, and following or stalking were included in the Controlling Abuse Scale, but making threats to leave, insulting, or shaming a partner were not included. The selection of controlling items was based on face validity, and it paralleled items in Ehrensaft et al. (1999) that showed high interrater agreement on a Q-sort task by a team of psychiatrists and clinical psychologists. The α reliability was .76 for study members' reports of perpetration and .82 for their reports of victimization on this scale. The physical abuse and controlling abuse reports were significantly correlated, indicating that men and women involved in one type of abuse were also involved in the other type of abuse ($r = .43$ for perpetration, $r = .57$ for victimization, ps $< .05$).

Designating Clinically Abusive, Nonclinically Abusive, and Nonabusive Relationships

We classified study members' relationships into one of three mutually exclusive types: nonabusive ($n = 746$), nonclinically abusive ($n = 134$), or clinically abusive ($n = 75$).

Members of the clinical abuse group were study members who reported violence resulting in one or more of the following reported on the PCC: injury (sprains, bruises, cuts, knocked out/loss of consciousness, broken bones, loose teeth), need for medical treatment (first aid, emergency room, overnight hospitalization), police intervention, and/or help-seeking for abuse (from a woman's shelter, a marriage therapist, a lawyer, or the courts via restraining orders). Clinical abuse was also defined if there was evidence of a court conviction for such behavior in the computerized police records of New Zealand or Australia (charges related to abuse in this cohort were "male assaults female, with weapon," "common manual assault, domestic" and "breeches nonmolestation order"). We obtained records for study members but not partners because partners did not take part in the age-26 assessments, and thus we could not obtain informed consent for a police record search from partners. It is clear from the New Zealand police records that the domestic assaults were against partners, not household members such as parents or children. The number of birth cohort members having a court conviction for partner assault is presented in Table 1: 0 women and 6 men. This constitutes a rate of 1% of men.

The nonclinical abuse group was composed of study members who had endorsed items on the Physical Abuse scales but reported none of the above-mentioned consequences on the PCC and did not have a conviction record.

The term clinical abuse is intended to be descriptive. It signifies that the cases so labeled had some type of clinical consequences, in terms of injury or receipt of services from agencies who deliver services in abuse cases. We defined the cases this way to test the hypothesis that cases identified on the basis of clinical criteria differ substantively from cases identified through scales such as the CTS in community samples. The clinical abuse group consisted of individuals who received services or who were injured, a group purported to be distinctive from those who report abuse on scales but do not have clinical consequences.

Table 1
Percentage of Men (n = 37) and Women (n = 38) in Clinically Abusive Relationships Endorsing Each Abuse Consequence on the Partner Conflict Calendar

	Female reporter		Male reporter	
Abuse consequence	%	*n*	%	*n*
Individual has domestic violence conviction record[a]	0_b		15_c	6
Police ever called	24	9	24	9
Woman injured	68	26	35	13
Man injured	26	10	60	22
Self needed medical care	24_d	9	3_e	1
Partner needed medical care	3	1	14	5
Woman sought help	34	13	38	14
Man sought help	24	9	32	12
Woman used shelter	3	1	0	
Man used shelter	0		0	

Note. Difference between subscripts b and c is significant at $p < .01$; difference between subscripts d and e is significant at $p < .05$.
[a] Official record.

Results

Description of the Clinically Abusive Relationships

Nine percent of the sample (women, $n = 38$; men, $n = 37$) reported being involved in a clinically abusive relationship. Table 1 shows the percentage (and number) of the clinical abuse group who endorsed each of the PCC items that could qualify a study member's relationship as clinically abusive. Men and women both reported high levels of injury to themselves (68% of women and 60% of men in the clinically abusive relationships said they were injured by their partner). Men and women tended to seek help from services (lawyers, marriage therapists, or the courts) at similar rates. Of particular note, men in clinically abusive relationships were more likely than women to have an official domestic violence conviction record (0% of women vs. 15% of men, $p < .01$), and women reporting on their abusive relationships were significantly more likely than men reporting on their abusive relationships to have needed medical care (24% of women vs. 3% of men, $p < .05$). No men reported that their partners had used a shelter and 1 woman in the clinical abuse group (3%) reported using a shelter; women's refuges are available throughout New Zealand.

Comparing Abuse Patterns in Clinically and Nonclinically Abusive Relationships

We examined the chronicity (mean number of months with partner violence), frequency (mean number of incidents per violent month), presence of alcohol or drugs (mean percentage of incidents with alcohol or drugs per violent month), and reports of the respondent's and his or her partner's controlling abuse in the two abuse groups. As summarized in Table 2, we report the responses of male reporters and female reporters separately on each of these three variables. Across both sexes, relative to the nonclinical abuse group, the clinical abuse groups experienced abuse that lasted longer on average (4.5 months vs. 3.0 months), $t(208) = 2.10$, $p < .05$, and was more frequent (average rate of three incidents vs. one incident of violence each month), $t(208) = 3.80$, $p < .001$. Clinical abuse couple members (one or both) were intoxicated by drugs or alcohol in approximately 50% of the incidents (ranging from none to all of the incidents) compared with 35% in the nonclinical group; this group difference approached significance, $t(208) = 1.66$, $p = .06$. Last, Table 2 shows that women in the clinical abuse group reported receiving more controlling abuse acts from their partners than women in the nonclinical abuse group, $t(101) = 2.71$, $p < .01$; for men, this group difference was not significant, $t(104) = 1.61$, $p = .11$. Neither men nor women differed across the two abuse groups on their reports of their own controlling abuse perpetration.

Notably, the clinical abuse groups experienced more months with abuse and more incidents despite having less exposure time, because their relationships were of shorter duration ($M = 33.4$ months, $SD = 33.6$ months) compared with the nonclinical groups' relationships ($M = 47.6$ months, $SD = 31.3$ months), $t(204) = 3.04$, $p < .01$.

Table 2
Mean Chronicity and Frequency of Substance Use During Partner Abuse and Partner Controlling Abuse for Nonclinically and Clinically Abusive Relationships

	Male reporter: type of abusive relationship		Female reporter: type of abusive relationship	
Scale	Nonclinical	Clinical	Nonclinical	Clinical
Partner Conflict Calendar				
Months with partner abuse (range)	$3.2\ (1–31)_a$	$4.8\ (1–32)_b$	$2.8\ (1–37)_a$	$4.5\ (1–21)_b$
Incidents per abusive month (range)	$1.2\ (1–4.5)_a$	$1.8\ (1–8)_b$	$1.2\ (1–4)_a$	$3.9\ (1–30.5)_b$
Percentage of incidents with alcohol or drugs per abusive month (range)	39 (0–100)	43 (0–100)	$33\ (0–100)_c$	$51\ (0–100)_d$
Controlling Abuse Scale[a]				
Partner's controlling abuse (*SD*)	$.40\ (1.12)_c$	$.81\ (1.50)_d$	$.21\ (1.02)_a$	$.97\ (1.80)_b$
Own controlling abuse (*SD*)	.50 (1.18)	.91 (1.87)	.37 (1.15)	.33 (1.11)
n	69	38	65	37

Note. Means with subscripts a and b differ significantly from one another at $p < .05$ in *t* tests. Means with subscripts c and d differ marginally from one another at $p = .10$.
[a] Mean scale scores standardized to a *z*-score metric.

To obtain a clearer clinical impression of the two types of abusive relationships, we compared men and women in these relationships on levels of physical abuse toward romantic partners, using our standardized Dunedin Abuse scales. We conducted focused comparisons (*t* tests), rather than omnibus analyses, to specifically compare the two abuse groups across the sexes and compare the sexes within each abuse group. This focused data-analytic approach is advocated by Rosenthal, Rosnow, and Rubin (2000).

Figure 1 shows the means on the Physical Abuse scale, by relationship type and sex of the reporter. First, we compared the sexes, within relationship type. In nonclinical abusive relationships ($n = 134$), women reported perpetrating significantly more physical abuse than did men, $t(132) = 3.27$, $p < .001$, and being victimized by significantly less physical abuse than men, $t(132) = 5.23$, $p < .001$. In contrast, in clinically abusive relationships ($n = 75$), men and women did not differ from each other in their reports of perpetration, $t(73) = 0.19$, $p > .10$, or victimization, $t(73) = 0.25$, $p > .10$. Second, we compared relationship types, within sex. Women in clinically abusive relationships did not differ in their reports of perpetration from women in nonclinically abusive relationships, $t(101) = -0.88$, $p > .10$, but they did report being victimized by significantly more physical abuse than did women in nonclinically abusive relationships, $t(101) = -4.73$, $p < .001$. Consistent with the women's view, the men in clinically abusive relationships reported perpetrating significantly more physical abuse, $t(104) = -2.28$, $p < .001$, and they also reported being victimized by more physical abuse than did men in nonclinically abusive relationships, $t(104) = -2.08$, $p < .001$.

Developmental Risk Factors for Abusive and Nonabusive Relationships

Next, we tested developmental risk factors for becoming involved in any type of abusive relationship and whether these risks discriminated between becoming involved in a clinically versus nonclinically abusive relationship. The raw scores for childhood developmental variables were first standardized on the entire cohort, by sex, using the *z*-score transformation. Tables and figures show the standardized scores ($M = 0$, $SD = 1$) for the developmental variables in each type of young-adult relationship, for women and men, respectively.

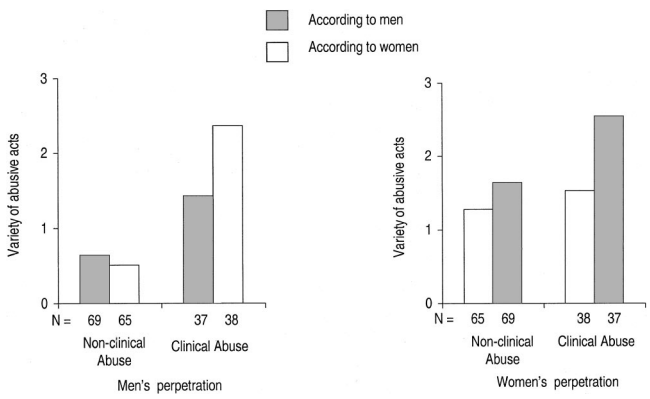

Figure 1. Means on physical abuse scale by relationship type and sex according to men's reports and women's reports.

We tested planned contrasts between the following pairs of relationship types: nonabusive (NA) versus nonclinically abusive (NCA); nonabusive (NA) versus clinically abusive (CA); and nonclinically abusive (NCA) versus clinically abusive (CA). The planned comparisons were carried out in a regression framework, in which we used dummy variables to represent the groups being compared on each of the developmental risk factors. This allowed us to conduct ordinary least squares (OLS) regression analyses when the developmental risk factor was a continuous variable (with corresponding *t* tests that were equivalent to planned contrasts between each of the abuse groups in an analysis of variance [ANOVA] framework) and logistic regressions when the developmental risk factor was categorical (which could not be done in an ANOVA framework given the 0/1 nature of the risk factor).

The prospective risk factor profiles were somewhat different for men and women in abusive relationships. As shown in Table 3, women who became involved in abusive relationships as adults experienced more caretaker changes in childhood, spent more years with a single parent, and were more likely to have a juvenile diagnosis of *conduct disorder* between ages 11 and 15 years than women who did not become involved in abusive relationships. Planned contrasts suggested that being reared by a single parent and conduct disorder were risk factors for women who became involved in both nonclinically and clinically abusive relationships.

Table 4 shows that men who became involved in abusive relationships as adults, unlike women, were indistinguishable on the basis of their family-of-origin characteristics, except for higher rates of harsh discipline. However, a history of externalizing childhood and adolescent behavior problems (as indexed by parents' and teachers' reports) and adolescent clinical diagnoses of conduct disorder and attention deficit disorder distinguished the men in clinically abusive relationships. Planned contrasts indicated that the risk factors applied mainly to the men in clinically abusive relationships; men in nonclinically abusive relationships were not distinguishable from those in nonabusive relationships by childhood risk factors, except on hyperactivity.

Adolescent Personality Profiles

Last, we tested whether adolescent personality differences discriminated between women and men in each of the three types of adult relationships (Figure 2 for men, Figure 3 for women). We conducted a multivariate ANOVA (MANOVA) on the 10 personality scales, within sex, contrasting the groups as follows: nonabusive to nonclinical abuse, nonabusive to clinical abuse, and nonclinical abuse to clinical abuse. For women, the multivariate test was not significant (Wilks's $\lambda = .95$), $F(20, 886) = 1.24$, $p > .10$. Univariate tests suggested that the women differed only on the Aggression scale, $F(2, 452) = 6.78$, $p < .001$; specifically, those in the nonclinical abusive group scored higher than those in the nonabusive group ($d = .32$, $p < .05$), and women in the clinical abuse group scored higher than those in the nonabusive group ($d = .52$, $p = .01$). In contrast, for men, the multivariate test was significant (Wilks's $\lambda = .88$), $F(20, 904) = 3.07$, $p < .001$. Univariate tests indicated that men in abusive relationships were statistically deviant on half of the personality scales, including low Traditionalism, $F(2, 461) = 7.08$, $p < .001$; low Self-Control, $F(2,$

Table 3
Developmental Risk Factors for Women in Nonabusive (NA), Nonclinically Abusive (NCA), and Clinically Abusive (CA) Relationships

Developmental risk factors	Type of relationship (*SD*)			Planned contrast: *t* statistic (effect size)		
	NA (n = 373)	NCA (n = 67)	CA (n = 38)	NA vs. NCA	NA vs. CA	NCA vs. CA
Family-of-origin characteristics						
SES (average from Ages 1 to 15)	.02 (.05)	.10 (.12)	−.23 (.16)	−0.54 (0.01)	1.49 (0.21)	1.60 (0.22)
No. of caretaker changes (to Age 11)	−.06 (.05)	.16 (.13)	.37 (.17)	−1.63 (0.22)	−2.38 (0.43)*	−0.97 (0.21)
Years with single parent (to Age 11)	−.08 (.05)	.25 (.12)	.25 (.16)	−2.55 (0.33)**	−2.02 (0.33)*	0.003 (0.001)
Mother's age at first birth	.03 (.05)	−.14 (.13)	.07 (.17)	1.28 (0.17)	−0.24 (0.04)	−1.02 (0.22)
Parenting						
Mother's rejecting behavior, Age 3	−.03 (.05)	.19 (.12)	−.17 (.16)	−1.68 (0.22)	0.82 (0.14)	1.78 (0.36)
Inconsistent parenting, Ages 7 and 9	−.01 (.05)	.04 (.13)	.06 (.17)	−0.36 (0.05)	−0.43 (0.08)	−0.12 (0.02)
Harsh discipline, Ages 7 and 9	.02 (.05)	.01 (.13)	−.29 (.17)	0.10 (0.01)	1.43 (0.31)	1.78 (0.30)
Child behavior problems						
Teacher-rated antisocial, average age 5–11	−.004 (.05)	−.04 (.12)	.12 (.16)	0.58 (0.08)	−0.02 (0.002)	−0.40 (0.08)
Parent-rated antisocial, average age 5–11	.002 (.05)	.04 (.12)	−.004 (.16)	−0.27 (0.04)	−0.04 (0.006)	0.21 (0.04)
Teacher-rated hyperactive, average age 5–11	−.01 (.05)	−.10 (.12)	.06 (.16)	0.82 (0.11)	−0.30 (0.05)	−0.79 (0.16)
Parent-rated hyperactive, average age 5–11	.03 (.05)	−.13 (.12)	.003 (.16)	1.21 (0.16)	0.15 (0.03)	0.67 (0.14)
Child and adolescent diagnoses				Odds ratio (effect size[a]) (confidence interval)		
Conduct disorder, Age 11, 13, or 15	11%	23%	30%	2.36* (0.47) (1.21–4.54)	3.37** (0.67) (1.55–7.33)	1.44 (0.20) (0.58–3.57)
Attention deficit disorder, Age 11, 13, 15	3%	2%	3%	0.54 (0.34) (0.07–4.27)	1.00 (0) (0.12–8.02)	1.86 (0.34) (0.11–30.60)
Major depression, Age 11, 13, or 15	7%	6%	8%	0.83 (0.10) (0.28–2.45)	1.16 (0.08) (0.33–4.05)	1.41 (0.18) (0.30–6.67)

Note. Continuous variables are represented by z scores (M = 0, SD = 1) and categorical variables by percentages. SES = socioeconomic status.
[a] To calculate effect sizes, we used the method described by Haddock, Rindscopf, and Shaddish (1998).
* $p < .05$. ** $p < .01$.

Table 4
Developmental Risk Factors for Men in Nonabusive (NA), Nonclinically Abusive (NCA), and Clinically Abusive (CA) Relationships

	Type of relationship (*SD*)			Planned contrast: *t* statistic (effect size)		
Developmental risk factors	NA (n = 373)	NCA (n = 67)	CA (n = 37)	NA vs. NCA	NA vs. CA	NCA vs. CA
Family-of-origin characteristics						
SES (average from Ages 1 to 15)	.02 (.05)	−.04 (.12)	−.15 (.16)	0.56 (0.06)	1.07 (0.17)	0.58 (0.11)
No. of caretaker changes (to Age 11)	.01 (.05)	−.03 (.13)	.004 (.17)	0.26 (0.04)	0.02 (0.004)	−0.15 (0.03)
Years with single parent (to Age 11)	−.02 (.05)	.03 (.12)	.01 (.16)	−0.38 (0.05)	−0.18 (0.03)	0.10 (0.10)
Mother's age at first birth	.04 (.05)	−.21 (.13)	−.08 (.17)	1.88 (0.25)	0.72 (0.12)	−0.65 (0.13)
Parenting						
Mother's rejecting behavior, Age 3	−.02 (.05)	−.02 (.12)	.03 (.16)	−0.05 (0.07)	−0.32 (0.05)	−0.23 (0.04)
Inconsistent parenting, Ages 7 and 9	.02 (.05)	−.06 (.12)	.02 (.16)	0.58 (0.08)	−0.01 (0.002)	−0.40 (0.08)
Harsh discipline, Ages 7 and 9	−.05 (.05)	−.03 (.12)	.28 (.16)	−0.15 (0.02)	−1.98 (0.33)*	−1.55 (0.31)
Child behavior problems						
Teacher-rated antisocial, average age 5–11	−.08 (.05)	.16 (.12)	.33 (.15)	−1.82 (0.23)	−2.55 (0.41)**	−0.92 (0.18)
Parent-rated antisocial, average age 5–11	−.05 (.05)	.16 (.12)	.08 (.15)	−1.67 (0.22)	−0.79 (0.13)	0.45 (0.08)
Teacher-rated hyperactive, average age 5–11	−.10 (.05)	.23 (.12)	.32 (.15)	−2.66 (0.34)**	−2.67 (0.42)**	−0.46 (0.08)
Parent-rated hyperactive, average age 5–11	−.07 (.05)	.11 (.12)	.31 (.15)	−1.39 (0.18)	−2.34 (0.38)*	−1.04 (0.20)
Child and adolescent diagnoses				Odds ratio (effect size[a]) (confidence interval)		
Conduct disorder, Age 11, 13, or 15	25%	33%	54%	1.48 (0.22) (0.84–2.62)	2.66** (0.53) (1.37–5.15)	1.79 (0.32) (0.80–3.99)
Attention deficit disorder, Age 11, 13, 15	7%	13%	24%	2.03 (0.39) (0.91–4.55)	4.22** (0.79) (1.86–9.54)	2.08 (0.40) (0.76–5.65)
Major depression, Age 11, 13, or 15	4%	3%	10%	0.81 (0.12) (0.23–2.81)	1.87 (0.34) (0.61–5.76)	2.31 (0.46) (0.49–10.87)

Note. Continuous variables are represented by *z* scores (M = 0, SD = 1) and categorical variables by percentages. SES = socioeconomic status.
[a] To calculate effect sizes, we used the method described by Haddock, Rindscopf, and Shaddish (1998).
* $p < .05$. ** $p < .01$.

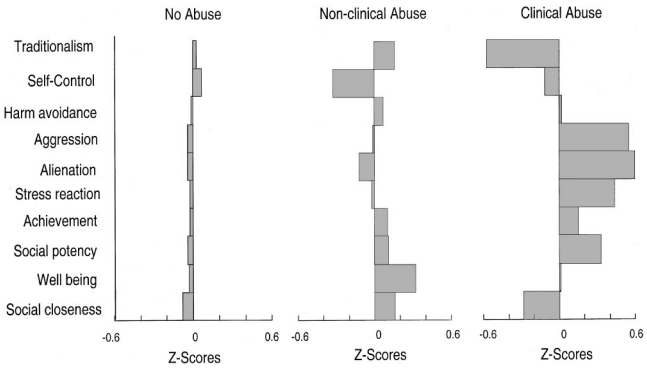

Figure 2. Adolescent personality profiles (age = 18 years) of men who later became involved in nonabusive or nonclinically abusive versus clinically abusive relationships.

461) = 4.34, $p < .01$; high Alienation, $F(2, 461) = 7.68$, $p < .001$; high Aggression, $F(2, 461) = 6.15$, $p < .01$; and high Stress Reaction, $F(2, 461) = 3.69$, $p < .05$. Specifically, men in nonclinically abusive relationships scored lower than men in nonabusive relationships on Self-Control ($d = .39$, $p < .01$). Men in clinically abusive relationships scored higher than men in nonclinically abusive relationships on Stress Reaction ($d = .46$, $p < .05$), Alienation ($d = .72$, $p < .001$), and Aggression ($d = –.57$, $p < .01$), and lower on Traditionalism ($d = .73$, $p < .001$) and Social Closeness ($d = .44$, $p < .05$; although the omnibus univariate test was not significant for this variable, the difference between these two abuse groups was significant).

Discussion

This study demonstrated that partner abuse that has clinical consequences can be identified in a community sample and that the epidemiological characteristics, correlates, and predictors of such abuse differ from those of nonclinical abuse. The prevalence of clinically significant partner violence, resulting in injury or official intervention, in the general population of young adults is 9%, but this prevalence rate requires replication.

Recent surveys of representative community samples have measured the prevalence of frequent (Capaldi & Owen, 2001) or "severe" partner aggression (Kessler, Molnar, Feurer, & Appelbaum, 2001; MacMillan & Gartner, 1999; Schaefer, Caetano, & Clark, 1998) and associated injury (Sorenson, Upchurch, & Shen, 1996; Stets & Straus, 1990). These previous findings are limited by inclusion of only respondents who were in current relationships at the time of the interview and by the absence of data on other consequences of the partner violence. Our results extend these earlier findings about severity, to arrive at a more comprehensive definition of abuse that would be considered clinically significant by health professionals. The clinically abusive group identified in this complete birth cohort contains both cases of abusers and victims who are typical in court-mandated treatments or women's shelters, plus other cases not seen in courts or shelters but that are nonetheless significant from public health and clinical points of view.

The present classification of abuse with the PCC advances previous work in its simultaneous inclusion of injury, requiring (with or without receiving) medical attention, seeking legal or police intervention, or using a shelter. In fact, one third of men and women in such relationships would have been missed by asking exclusively about injuries. The PCC measure of abuse extends the period of assessment to 3 years (rather than the 12 months that are typical of partner violence research) without compromising reliability. Furthermore, the study measured abuse in both intact and dissolved relationships and collected data about the context of partner abuse, such as partners' use of alcohol, frequency, abuse, et cetera. Last, the measure was administered to both men and women.

Prior literature suggested that (a) *common couples abuse* (the less severe form of abuse) would characterize mutual abuse practiced by both men and women and (b) clinical abuse (the form severe enough to produce injury and agency intervention) would not be mutual but would involve mainly man-to-woman abuse. We found something unexpected. The less severe form involved primarily woman-to-man abuse, but the clinical form involved abuse practiced by both men and women.

We have reason for confidence in these findings because both male and female reporters agreed about these patterns, this cohort has developed trust in the confidentiality guarantee over repeated assessments, and study members' reports were corroborated previously by their partners (Moffitt et al., 1997). Independent findings demonstrated that young men and women in community samples of couples reported similarly high levels of injuries (Capaldi & Owen, 2001).

Women in nonclinically and clinically abusive relationships were similar in that both groups had aggressive personalities and/or adolescent conduct disorder and reported using similar levels of abusive behaviors. Men in nonclinically abusive relationships were similar to nonabusive men; they had no notable developmental antecedents, no personality deviance, and as a group they used very few abusive behaviors (suggesting the possibility that their female partners' greater levels of abuse were not in self-defense). Men in clinically abusive relationships were markedly different; they had psychopathology both in childhood and adolescence, extensive personality deviance, and used a wide variety of abusive behaviors against their partners. What distinguishes abusive relationships that become clinically significant from those that do not is that, in clinically abusive relationships, men perpe-

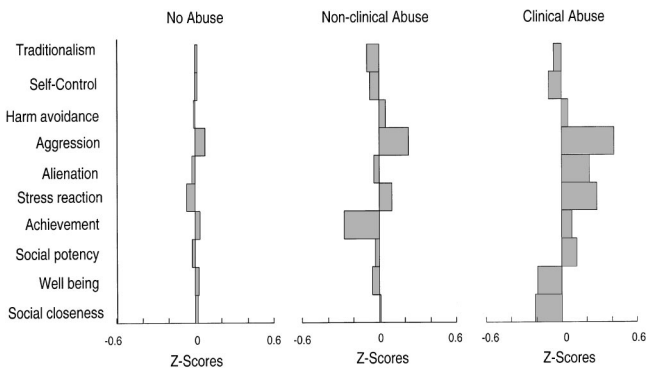

Figure 3. Adolescent personality profiles (age = 18 years) of women who later became involved in nonabusive or nonclinically abusive versus clinically abusive relationships.

trate abuse (and so do women). When men take an active part in abuse, abuse lasts more months with more incidents per month, more men are convicted of domestic violence, and more women need medical treatment for injuries. When only the woman takes part in abuse, such clinical consequences do not ensue. Thus, the participation of men in clinically abusive relationships is associated with worse health consequences for women and greater health-delivery costs (for review, see Campbell, 2002). Women, but not men, in the clinical abuse group also reported higher levels than those in the nonclinical abuse group of controlling abuse from their partner.

These findings have implications for theories of partner abuse. First, our findings are not consistent with the theory of patriarchal societal norms as the main cause of clinical partner abuse (Johnson, 1995; MacMillan & Gartner, 1999). This theory argues that serious partner abuse is not driven primarily by a man's personal pathology; rather, it is motivated by societal acceptance of gender-based inequality and violence against women. Some proponents of this theory are beginning to change their views on the contribution of male psychopathology to serious partner abuse (Johnson & Ferraro, 2000). Our results support such a change in views; men who are clinically abusive exhibited long-standing disinhibitory behavioral pathology. These men were especially likely to score low on the Traditionalism scale of our personality assessment, which is inconsistent with the notion that violence against women is motivated by conventional, normative patriarchal attitudes. Second, our findings support the assertion of advocates that female victims are not more pathological than other women (Dobash & Dobash, 1992; Walker, 1989). However, the data show that they are more generally aggressive. Third, clinical abuse seems to require a pathologically abusive man. This prompts the novel hypothesis that woman-to-man abuse is the common default, but escalation beyond this common pattern, to a more severe level involving injuries and official intervention, requires a male partner who has a history of psychopathology. This hypothesis requires a direct test and is indicated by our earlier findings that both partners' risk profiles additively and independently increase the risk for partner abuse (Moffitt, Robins, et al., 2001).

The finding that both men and women in clinically abusive relationships tend to have a history of aggression agrees with evidence of assortative mating between antisocial individuals (Galbaud du Fort, Bland, Newman, & Boothroyd, 1998; Krueger, Moffitt, Caspi, & Bleske, 1998) and with studies of community samples that have found that antisocial behavior is associated prospectively with partner violence (Andrews et al., 2000; Capaldi & Owen, 2001; Ehrensaft, Cohen, et al., 2003). Women in clinically abusive relationships show a history of conduct problems that begin in adolescence; this adolescent-onset pattern is typical of the course of women's antisocial behavior reported elsewhere (for reviews, see Moffitt, Caspi, et al., 2001; Silverthorn & Frick, 1999). This suggests the developmental hypothesis that one factor in the origins of clinically abusive relationships is the assortative mating of antisocial men who have long-standing histories of behavioral disinhibition, with women who begin to display serious behavior problems in adolescence. Thus, one of the "snares" for girls who develop conduct problems may be involvement with abusive antisocial males.

These theoretical considerations also have implications for preventing partner violence. First, male and female adolescents with a current or previous history of antisocial behavior should be targeted for prevention services for their own health and to prevent exposure of future generations to partner abuse (Moffitt, Caspi, et al., 2001). Exposure to domestic violence damages children's intellectual and behavioral development, even net of genetic transmission in the family and other co-occurring risk processes (Koenen, Moffitt, Caspi, Taylor, & Purcell, 2003; Yates, Dodds, Sroufe, & Egeland, 2003). At present, there is a disconnect between services for juvenile delinquency and services for partner abuse; our findings of significant overlap call for an integration of these two types of services. Second, prevention programs and treatments must address both men's and women's aggression toward partners. Traditionally, intervention and prevention programs have exclusively targeted male partner abuse. This study and others (Capaldi & Owen, 2001; Moffitt, Robins, et al., 2001) suggest that this single-sex approach is not empirically supported, because both partners' behaviors contribute to the risk of clinically significant partner abuse, and both partners should be treated. Women's partner abuse cannot be explained exclusively as self-defense against men's partner abuse, because a woman's pre-relationship history of aggression toward others predicts her abuse toward her partner, over and above controls for her reports of his abuse toward her (Moffitt, Robins, et al., 2001). Third, interventions for partner violence, which currently have limited effectiveness (Dunford, 2000), might be strengthened by borrowing concepts from successful delinquency interventions (e.g., Borduin et al., 1995; Chamberlain & Reid, 1998), which focus more broadly on aggressive, antisocial behavior in boys and girls, rather than on changing men's attitudes and behavior toward women.

Finally, our findings have implications for research methods and sample selection in future research studies. A major controversy concerns whether findings from epidemiological samples apply to clinical abuse. This study finds that clinical abuse resulting in serious injury or official intervention can indeed be ascertained without bias by using an epidemiological sample. Our findings argue against the suggestion that shelter, emergency room, and police samples necessarily tap a categorically different (and more consequential) type of abuse than community samples (Johnson, 1995). Because more women are clinically affected by serious partner abuse than the selective minority in battered women's shelters, police precincts, or emergency rooms (Straus, 1999), we would argue for continued development of sensitive and specific instruments to measure clinical partner abuse in representative community samples. These instruments could be used in primary care settings to identify persons in need of intervention (Wiist & McFarlane, 1999).

This study has several important limitations. First, the PCC measures of abuse consequences were based on a single informant reporting on the relationship. Thus, some cases of clinical partner abuse may have been missed because of partners who did not know whether their partner was injured or sought official intervention. If so, then some clinically significant cases may have been counted among the nonclinical cases, reducing the chance to observe the group differences we found. Second, our data covered a 3-year period, and the participants may have been subject to forgetting. However, the reliability of 3-year recall on the PCC was excellent, and, regardless, there is no reason to expect sex differences in recall accuracy that would make our sex comparisons artifactual. Moreover, if individuals who forgot abuse or its

consequences were assigned to groups incorrectly, this would have reduced the chances of detecting group differences. Third, although we used a reliable calendar-based instrument to measure partner violence over a 3-year period in men and women in their late 20s, this age group is at high risk for partner violence (U.S. Department of Justice, 1995), so the prevalence rate obtained here may not generalize to older ages. However, it seems reasonable to study partner abuse in the young age group, who suffers the primary burden among victims and their young children. Fourth, we lacked adequate power to detect small differences on developmental risk factors for the two abuse groups, or Group × Sex Difference interactions within the abuse groups. The differing risk profiles for men and women in the two groups are to be interpreted, therefore, with caution and require replication with larger samples to test for significant differences. Fifth, although clinically abusive relationships were uniquely characterized by bidirectional abuse, the present findings shed no light on the proximal processes by which a clinically abusive relationship develops over time. Future process research is needed to explain men's and women's experience of partner abuse over time, both within and across relationships. This study implies that such research is best conducted in the context of a sample that represents clinical abuse in an unbiased fashion.

References

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

Andrews, J. A., Foster, S. L., Capaldi, D., & Hops, H. (2000). Adolescent and family predictors of physical aggression, communication and satisfaction in young adult couples: A prospective analysis. *Journal of Consulting and Clinical Psychology, 68,* 195–208.

Archer, J. (2000). Sex differences in aggression between heterosexual partners: A meta-analytic review. *Psychological Bulletin, 126,* 651–680.

Bachman, R., & Saltzman, L. E. (1995). *Violence against women: Estimates from the redesigned survey* (BJS Publication No. 154-348). Washington, DC: Bureau of Justice Statistics, U.S. Department of Justice.

Belli, R. F., Shay, W. L., & Stafford, F. P. (2001). Event history calendars and question list surveys: A direct comparison of interviewing methods. *Public Opinion Quarterly, 65,* 45–74.

Borduin, C. M., Mann, B. J., Cone, L. T., Henggeler, S. W., Fucci, B. R., Blaske, D. M., & Williams, R. A. (1995). Multisystemic treatment of serious juvenile offenders: Long-term prevention of criminality and violence. *Journal of Consulting and Clinical Psychology, 63,* 569–578.

Buve, A., Carael, M., Hayes, R. J., Auvert, B., Ferry, B., Robinson, N. J., et al. (2001). Multicentre study on factors determining differences in rate of spread of HIV in sub-Saharan Africa: Methods and prevalence of HIV infection. *AIDS, 15*(Suppl. 4), S5–S14.

Campbell, J. (2002). Health consequences of intimate partner violence. *Lancet, 359,* 1331–1336.

Capaldi, D. M., & Clark, S. (1998). Prospective family predictors of aggression toward female partners for at-risk young men. *Developmental Psychology, 34,* 1175–1188.

Capaldi, D., & Owen, L. D. (2001). Physical aggression in a community sample of at-risk young couples: Gender comparisons for high frequency, injury and fear. *Journal of Family Psychology, 15,* 425–440.

Caspi, A., Moffitt, T. E., Thornton, A., Freedman, D., Amell, J. W., Harrington, H., Smeijers, J., & Silva, P. A (1996). The life history calendar: A research and clinical assessment method for collecting retrospective event-history data. *International Journal of Methods in Psychiatric Research, 6,* 101–114.

Chamberlain, P., & Reid, J. B. (1998). Comparison of two community alternatives to incarceration for chronic juvenile offenders. *Journal of Consulting and Clinical Psychology, 66,* 624–633.

Church, T. A., & Burke, P. J. (1994). Exploratory and confirmatory tests of the Big Five and Tellegen's three- and four-dimensional models. *Journal of Personality and Social Psychology, 66,* 93–114.

Costello, A., Edelbrock, C., Kalas, R., Kessler, M., & Klaric, S. (1982). *National Institute of Mental Health Diagnostic Interview Schedule for Children.* Rockville, MD: National Institute of Mental Health.

Dobash, R. E., & Dobash, R. P. (1992). *Women, violence and social change.* London: Routledge.

Dunford, F. W. (2000). The San Diego Navy experiment: An assessment of interventions for men who assault their wives. *Journal of Consulting and Clinical Psychology, 68,* 468–476.

Ehrensaft, M. K., Cohen, P., Brown, J., Smailes, E., Chen, H., & Johnson, J. (2003). Intergenerational transmission of partner violence: A 20-year prospective study. *Journal of Consulting and Clinical Psychology, 71,* 741–753.

Ehrensaft, M. K., Langhinrichsen-Rohling, J., Heyman, R. E., O'Leary, K. D., & Lawrence, E. (1999). Feeling controlled in marriage: A phenomenon specific to physically aggressive couples? *Journal of Family Psychology, 13,* 20–32.

Ehrensaft, M. K., & Vivian, D. (1999). Is physical aggression between dating partners related to their appraisals of coercive control? *Journal of Family Violence, 14,* 251–266.

Ehrensaft, M. K., Wasserman, G. A., Verdeli, H., Greenwald, S., Davies, M., & Miller, L. S. (2003). Maternal antisocial behavior, parenting practices, and behavior problems in boys at risk for antisocial behavior. *Journal of Child and Family Studies, 12,* 27–40.

Elander, J., & Rutter, M. (1996). Use and development of the Rutter parents' and teachers' scale. *International Journal of Methods in Psychiatric Research, 6,* 63–78.

Elley, W. B., & Irving, J. C. (1976). Revised socio-economic index for New Zealand. *New Zealand Journal of Educational Studies, 11,* 25–36.

Farrington, D. P., & West, D. J. (1990). The Cambridge study of delinquent development: A long-term follow-up of 411 London males. In H. J. Kerner & G. Kaiser (Eds.), *Kriminalitat* (pp. 117–138). New York: Springer-Verlag.

Frischer, M., Hickman, M., Kraus, L., Mariani, F., & Wiessin, L. (2001). A comparison of different methods for estimating the prevalence of problematic drug misuse in Great Britain. *Addiction, 96,* 1465–1476.

Galbaud du Fort, G., Bland, R. C., Newman, S. C., & Boothroyd, L. J. (1998). Spouse similarity for lifetime psychiatric history in the general population. *Psychological Medicine, 28,* 789–802.

Giordano, P. C., Millhollin, T. J., Cernkovich, S. A., Pugh, M. D., & Rudolph, J. L. (1999). Delinquency, identity and women's involvement in relationship violence. *Criminology, 27,* 17–40.

Gondolf, E. W. (1998). *Assessing woman battering in mental health services.* Thousand Oaks, CA: Sage Publications.

Haddock, C. K., Rindscopf, D., & Shaddish, W. R. (1998). Using odds ratios as effect sizes for meta-analysis: A primer on methods and issues. *Psychological Methods, 3,* 339–353.

Holtzworth-Munroe, A., Meehan, J. C., Herron, K., Rehman, U., & Stuart, G. L. (2000). Testing the Holtzworth-Munroe and Stuart (1994) batterer typology. *Journal of Consulting and Clinical Psychology, 69,* 1000–1019.

Holtzworth-Munroe, A., & Stuart, G. L. (1994). Typologies of male batterers: Three subtypes and the differences among them. *Psychological Bulletin, 116,* 476–497.

Hornung, C. A., McCullough, B. C., & Sugimoto, T. (1981). Status relationships in marriage: Risk factors in spouse abuse. *Journal of Marriage and the Family, 43,* 675–692.

Huesmann, L. R., Eron, L. D., Lefkowitz, M. M., & Walder, L. O. (1984). Stability of aggression over time and generations. *Developmental Psychology, 20,* 1120–1134.

Johnson, M. P. (1995). Patriarchal terrorism and common couple violence: Two forms of violence against women. *Journal of Marriage and the Family, 57,* 283–294.

Johnson, M. P., & Ferraro, K. J. (2000). Research on domestic violence in the 1990s: Making distinctions. *Journal of Marriage and the Family, 62,* 948–963.

Kessler, R. C., Molnar, B. E., Feurer, I., & Appelbaum, M. (2001). Patterns and mental health predictors of domestic violence in the United States: Results from the National Comorbidity Survey. *International Journal of Law and Psychiatry, 24,* 487–508.

Kim, H., & Capaldi, D. M. (in press). The association of antisocial behavior and depressive symptoms between partners and risk for aggression in romantic relationships. *Journal of Family Psychology*.

Koenen, K. C., Moffitt, T. E., Caspi, A., Taylor, A., & Purcell, S. (2003). Domestic violence is associated with environmental suppression of IQ in young children. *Development and Psychopathology, 15,* 297–311.

Koss, M. P., Goodman, L. A., Browne, A., Fitzgerald, L. F., Keita, G. P., & Russo, N. F. (1994). Understanding the perpetrator and the victim: Who abuses and who is abused? In M. P. Koss, L. A. Goodman, & A. Browne (Eds.), *No safe haven: Male violence against women at home, at work, and in the community* (pp. 19–38). Washington, DC: American Psychological Association.

Krueger, R. F., Caspi, A., & Moffitt, T. E. (2000). Epidemiological personology: The unifying role of personality in population-based research on problem behaviors. *Journal of Personality, 68,* 967–998.

Krueger, R. F., Moffitt, T. E., Caspi, A., & Bleske, A. (1998). Assortative mating for antisocial behavior: Developmental and methodological implications. *Behavior Genetics, 28,* 173–186.

Krueger, R. F., Schmutte, P. S., Caspi, A., & Moffitt, T. E. (1994). Personality traits are linked to crime among men and women: Evidence from a birth cohort. *Journal of Abnormal Psychology, 103,* 328–338.

Loseke, D. R. (1992). *The battered woman and shelters: The social construction of wife abuse*. Albany, NY: State University of New York Press.

MacMillan, R., & Gartner, R. (1999). When she brings home the bacon: Labor-force participation and the risk of spousal violence against women. *Journal of Marriage and the Family, 61,* 947–958.

Magdol, L., Moffitt, T. E., Caspi, A., & Silva, P. A. (1998). Developmental antecedents of partner abuse: A prospective-developmental study. *Journal of Abnormal Psychology, 107,* 375–389.

Moffitt, T. E., Caspi, A., Krueger, R. F., Magdol, L., Silva, P. A., & Sydney, R. (1997). Do partners agree about abuse in their relationships? A psychometric evaluation of interpartner agreement. *Psychological Assessment, 9,* 47–56.

Moffitt, T. E., Caspi, A., Rutter, M., & Silva, P. A. (2001). *Sex differences in antisocial behavior: Conduct disorder, delinquency, and violence in the Dunedin Longitudinal Study*. Cambridge, UK: Cambridge University Press.

Moffitt, T. E., Krueger, R. F., Caspi, A., & Fagan, J. (2000). Partner abuse and general crime: How are they the same? How are they different? *Criminology, 38,* 201–235.

Moffitt, T. E., Robins, R. W., & Caspi, A. (2001). A couples analysis of partner abuse with implications for abuse prevention. *Criminology and Public Policy, 1,* 5–36.

O'Leary, K. D. (1999). Developmental and affective issues in assessing and treating partner aggression. *Clinical Psychology: Science and Practice, 6,* 400–414.

O'Leary, K. D. (2000). Are women really more aggressive than men in intimate relationships? Comment on Archer (2000). *Psychological Bulletin, 126,* 685–689.

Patrick, C. J., Curtin, J. J., & Tellegen, A. (2002). Development and validation of a brief form of the Multidimensional Personality Questionnaire. *Psychological Assessment, 14,* 150–163.

Robins, L. (1978). Sturdy childhood predictors of adult antisocial behaviour: Replications from longitudinal studies. *Psychological Medicine, 8,* 611–622.

Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2000). Contrasts and effect sizes in behavioral research: A correlational approach. Cambridge, UK: Cambridge University Press.

Schaefer, J., Caetano, R., & Clark, C. L. (1998). Rates of intimate partner violence in the United States. *American Journal of Public Health, 88,* 1702–1704.

Silva, P. A., & Stanton, W. R. (1996). *From child to adult: The Dunedin Multidisciplinary Health and Development Study*. Auckland, New Zealand: Oxford University Press.

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

Sorenson, S. B., Upchurch, D. M., & Shen, H. (1996). Violence and injury in marital arguments: Risk patterns and gender differences. *American Journal of Public Health, 86,* 35–40.

Stets, J. E., & Straus, M. A. (1990). Gender differences in reporting marital violence and its medical and psychological consequences. In M. A. Straus & R. J. Gelles (Eds.), *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families* (pp. 151–166). New Brunswick, NJ: Transaction Publishers.

Straus, M. A. (1990). Measuring intrafamily conflict and violence: The Conflict Tactics (CT) Scales. In M. A. Straus & R. J. Gelles (Eds.), *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families* (pp. 403–424). New Brunswick, NJ: Transaction Press.

Straus, M. A. (1999). The controversy over domestic violence by women: A methodological, theoretical, and sociology of science analysis. In X. B. Arriaga & S. Oskamp (Eds.), *Violence in intimate relationships* (pp. 17–44). Thousand Oaks, CA: Sage.

Tjaden, P., & Thoennes, N. (1998). *Prevalence, incidence, and consequences of violence against women: Findings from the National Violence Against Women Survey*. Washington, DC: U.S. Department of Justice.

U.S. Department of Justice. (1995). *National Crime Victimization Survey: Violence against women. Estimates from the redesigned survey* (Special Report No. NCJ-154348). Washington, DC: Bureau of Justice Statistics.

Walker, L. E. (1989). *Terrifying love: Why battered women kill and how society responds*. New York: Harper & Row.

Wiist, W. H., & McFarlane, J. (1999). The effectiveness of an abuse assessment protocol in public health prenatal clinics. *American Journal of Public Health, 89,* 1217–1221.

Yates, T. M., Dodds, M. F., Sroufe, L. A., & Egeland, B. (2003). Exposure to partner violence and child behavior problems: A prospective study controlling for child physical abuse and neglect, child cognitive ability, socioeconomic status, and life stress. *Development & Psychopathology, 15,* 199–218.

Zlotnick, C., Kohn, R., Peterson, J., & Pearlstein, T. (1998). Partner physical victimization in a national sample of American families: Relationship to psychological functioning, psychosocial factors, and gender. *Journal of Interpersonal Violence, 13,* 156–166.

Received August 30, 2002
Revision received October 23, 2003
Accepted October 28, 2003 ■