Research Review: DSM-V conduct disorder: research needs for an evidence base

Terrie E. Moffitt,[1,2] and in alphabetical order: Louise Arseneault,[2]
Sara R. Jaffee,[2,3] Julia Kim-Cohen,[4] Karestan C. Koenen,[5] Candice L. Odgers,[2,6]
Wendy S. Slutske,[7] and Essi Viding[8]

[1]Duke University, USA; [2]King's College London, UK; [3]University of Pennsylvania, USA; [4]Yale University, USA;
[5]Harvard School of Public Health, USA; [6]University of California at Irvine, USA; [7]University of Missouri-Columbia,
USA; [8]University College London, UK

This article charts a strategic research course toward an empirical foundation for the diagnosis of conduct disorder in the forthcoming DSM-V. Since the DSM-IV appeared in 1994, an impressive amount of new information about conduct disorder has emerged. As a result of this new knowledge, reasonable rationales have been put forward for adding to the conduct disorder diagnostic protocol: a childhood-limited subtype, family psychiatric history, callous-unemotional traits, female-specific criteria, preschool-specific criteria, early substance use, and biomarkers from genetics, neuroimaging, and physiology research. This article reviews the evidence for these and other potential changes to the conduct disorder diagnosis. We report that although there is a great deal of exciting research into each of the topics, very little of it provides the precise sort of evidence base required to justify any alteration to the DSM-V. We outline specific research questions and study designs needed to build the lacking evidence base for or against proposed changes to DSM-V conduct disorder. **Keywords:** Conduct disorder, DSM-V.

This article charts a strategic research course toward an empirical foundation for the diagnosis of conduct disorder (CD) in the forthcoming fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-V). Since the publication of the DSM-IV (American Psychiatric Association, 1994), a great deal has been learned about CD. Now DSM-V is on the horizon. The DSM is a decision-making tool for clinicians, but also a framework for mental-health researchers. As such, there are hopes that DSM-V will bring innovations such as incorporating information about biological validity, dissolving disconnects between child and adult disorders, and including dimensional definitions as well as categorical diagnoses, among others (Phillips, First, & Pincus, 2005). It is time to reappraise DSM-IV in light of new information. It is also time to plan the remaining research priorities for the run-up to DSM-V. This article aims to inform this appraisal and planning process in two ways. First, the article presents a 'shopping list' of currently debated issues in the assessment and diagnosis of CD. We analyze 11 such issues as research priorities. Second, the article identifies research questions for each of the 11 issues, and suggests study designs to approach these questions. In so doing we aim to stimulate new research to improve the evidence base for decision-making by the future writers of the DSM-V section on CD.

This article is organized into four sections. The first section considers potential *changes* to the diagnosis of CD that are under consideration; including (1) adding CD subtypes, (2) incorporating family psychiatric history into the diagnosis of CD, and (3) adding psychopathic callous-unemotional traits as new criteria. The second section considers *biomarkers*. Hopes have been high that biomarkers will be ready to join the diagnostic criteria in time for DSM-V (Haag, 2007). Even if biomarkers are not ready to be diagnostic criteria, they will play a key role in DSM-V's aim to incorporate more biological information about construct validity. We evaluate the potential of biomarkers in the assessment of CD; including (4) neuroimaging, (5) genes, and (6) other neuro-physiological biomarkers such as heart rate, neurotransmitters, perinatal factors, and hormones. The article's third section considers current controversies about applying the diagnosis of CD to *subgroups* of young people; namely (7) preschool children and (8) girls. The fourth section considers perennial *conceptual issues* in the nosology of CD that remain relevant today for planning DSM-V. These conceptual issues include (9) the close association between CD and early substance use, (10) assessing CD as a category or a continuum, and (11) the life-course continuity from ODD (oppositional defiant disorder) to CD to ASPD (antisocial personality disorder). How did we identify these particular 11 issues? We began with issues suggested to us by the Committee to Assess Research Needs for DSM-V Externalizing Disorders. This article also draws on the meeting of this committee held in Mexico City in February, 2007. Other issues joined the list after each co-author sought recommendations from

---

clinicians and researchers in her sub-branch of the field. New issues will undoubtedly arise before DSM-V appears, but at present, these 11 seemed uppermost in the minds of CD professionals.

Our analysis of the aforementioned diagnostic issues follows certain guiding principles. Where possible we referenced reviews of the literature; if an authoritative review was not available to us, we cited key papers as examples instead of trying to include every relevant study. We considered each issue's predictive validity, assessment reliability, and rationales for and against incorporating it into the DSM-V diagnostic protocol for CD. We limited our focus to DSM-V, because regrettably we found little research into ICD-10 CD.

First, this article emphasizes the raison d'être of diagnosis: *predictive validity*. Our view is that even the most exciting research finding about CD will not be ready to inform DSM-V if its predictive validity is unknown. For each of the 11 issues, we ask if empirical evidence demonstrates that the information can improve clinicians' ability to predict patient outcomes. For example, can the information help to foretell CD patients' long-term prognosis, identify the subgroup of CD patients most in need of treatment, identify the most impaired or distressed CD patients, or predict CD patients' response to available treatments?

Second, this article emphasizes the pragmatics of diagnosis: *assessment reliability*. Even the most predictive research finding about CD will not be ready to inform DSM-V if it is too unwieldy, costly, or temporally unstable to be translated for use in clinical and forensic settings. For each issue, we asked if empirical evidence demonstrates that the information can be assessed reliably by clinicians at work in day-to-day practice. For example, are tools or protocols available that satisfy basic pragmatic criteria for test–retest reliability, inter-rater reliability, efficient use of clinician time, and affordable cost?

Third, as well as describing the rationale that supports considering each of the 11 diagnostic issues in discussions of DSM-V, this article articulates potential *disadvantages* of each issue. Typically the literature argues the case for adding a new criterion to DSM-V. In contrast, we noted that considering the case against a new criterion often generated a clearer picture of the unanswered questions, and pointed toward needs for research.

Issue 1, child-onset versus adolescent-onset subtypes of CD

Definition of the issue

DSM-IV distinguishes two subtypes of CD. The childhood-onset subtype (312.81) is defined by onset of at least one CD criterion prior to age 10 years, and the adolescent-onset type (312.82) is defined by the absence of any CD criterion prior to age 10 years.

DSM-IV notes that subtyping on the basis of age of onset captures differential information about the likely nature of a patient's characteristic presenting problems, developmental course, and prognosis (p. 86). The primary question for this section is: should this DSM-IV subtype system be kept? If age-of-onset subtyping is retained for DSM-V, must it be updated?

Rationale for reconsidering age-of-onset subtyping before DSM-V

Age-of-onset subtyping warrants reconsideration now for three reasons. First, at the time the DSM-IV was drafted, the evidence base to support subtyping on the basis of age of onset was rather sparse, consisting of just a few relatively short-term longitudinal studies (Lahey et al., 1998; Moffitt, 1993). Since then, the relevant literature has grown substantially, inviting us to revisit whether or not the evidence still supports the predictive validity of age-of-onset subtyping; does it really convey information about patients' characteristic problems, course, and prognosis, as the DSM-IV says? Second, we must query the reliability of the age-of-onset diagnosis. Can age of onset be assessed reliably in clinical as well as research settings? Third, we must query the utility of age-of-onset subtyping. Are clinicians using it? Does it generate useful information regarding treatment choice or treatment response? Challenging the status quo in this way is necessary because previous subtyping systems did not last. DSM-III laid out a matrix of four subtypes: socialized versus unsocialized, crossed with aggressive versus nonaggressive (American Psychiatric Association, 1980). DSM-IIIR presented three subtypes; solitary-aggressive, socialized, and undifferentiated with mixed features (American Psychiatric Association, 1987). These subtyping systems came and went rapidly for lack of predictive validity, assessment reliability, and clinical utility.

Predictive validity of age-of-onset subtypes

The evidence base provides good consensus for the distinction between a childhood-onset life-course persistent subtype and an adolescent-onset subtype, in both girls and boys. This evidence comes from longitudinal studies that have followed cohorts from childhood to adulthood and measured antisocial conduct at multiple time points. Many of these studies have now been analyzed using modern statistical methods that can detect subtype groups characterized by differential age of onset and long-term developmental course (Muthén & Muthén, 2004; Nagin, 2005). The resulting models typically uncover the expected large adolescent-onset group and a small childhood-onset group whose antisocial conduct persists for several years (Fergusson & Horwood, 2002; Lahey et al., 2006; Odgers et al.,

2007; Piquero, 2005; Wiesner, Kim, & Capaldi, 2005).

The evidence base confirms that these two subtypes convey differential information about patients' *characteristic problems*. The evidence comes from longitudinal studies of cohorts from over a dozen countries (Moffitt, 2003, 2006). Briefly, the childhood-onset persistent type is frequently characterized by severe family adversity, parental antisocial behavior and greater genetic liability, perinatal complications, neurocognitive deficits, low IQ, hyperactivity, inattention, impulsivity, school difficulties, and peer difficulties as children. The adolescent-onset type, by contrast, tends to score within normal limits on such problems. Research suggests that adolescent-onset CD youth are influenced by association with other delinquent youths, or by seeking social status through delinquent behaviors.

Evidence also indicates that these two subtypes predict patients' differential *course and prognosis*. Follow-ups to adulthood reveal relatively poorer adult outcomes for the childhood-onset persistent group in domains of violence, conviction, incarceration, personality disorder, other mental disorders, substance abuse, work life and family life (reviewed by Moffitt, 2003, 2006) as well as compromised physical health as shown by increased injuries, primary-care and hospital visits, sexually transmitted infections, systemic inflammation, periodontal disease, smoking, and chronic respiratory illness (Odgers et al., 2007; Odgers et al., in press; Piquero, Gibson, Daigle, Piquero, & Tibbetts, 2007).The adolescent-onset subtype fares relatively better. Their education, work, health and family-life outcomes are relatively unimpaired, but their adult prognosis includes substance abuse and crimes that go largely undetected (Odgers et al., 2007; Nagin, Farrington, & Moffitt, 1995). Thus both childhood-onset and adolescent-onset CD warrant intervention, but nevertheless it is worthwhile to diagnose the subtypes because they are thought to require different intervention goals and approaches (Howell & Hawkins, 1998; Scott & Grisso, 1997).

A possible childhood-limited developmental subtype

The evidence base suggests that a third subtype, not mentioned in DSM-IV, should also be considered here: *childhood-limited CD*. Robins (1966) first pointed out that one half of childhood-onset conduct-problem children do not grow up to have antisocial personality disorder. Longitudinal studies that track the continuity of conduct problems from childhood to adulthood have revealed the existence of an exceptional group of children who lack such continuity. These children are often termed the 'childhood-limited' conduct problem group (for a review see Moffitt, 2003, 2006). Some studies define this childhood-limited group broadly as a large group of

children having any elevated disruptive behavior; such studies remind us that mild, temporary conduct problems are ubiquitous in healthy young children. Studies of children with mild childhood-limited conduct problems report that such problems need not portend poor prognosis (Odgers et al., 2007; Tremblay, 2003). In contrast, other studies define this childhood-limited group narrowly as a small group of children exhibiting extreme, pervasive, and persistent conduct problems only during childhood. Follow-up studies indicate that these children with childhood-limited CD do not become antisocial adults, but they do become adults who are depressed, anxious, socially isolated, and financially dependent on others (Farrington, Gallagher, Morley, St. Ledger, & West 1988; Moffitt, Caspi, Harrington, & Milne 2002; Wiesner & Capaldi, 2003).

Reliability of assessing age-of-onset subtypes

Evidence suggests that clinicians are using the age-of-onset distinction; it is a key feature of manual-guided assessment protocols. The *Structured Assessment of Violence Risk in Youth* specifies 'early initiation' (SAVRY; Borum, Bartel, & Forth, 2006). The *Early Assessment Risk List for Boys* specifies first symptoms 'under age 6, versus age 7–12, versus over age 12' (EARL-20B; Augimeri, Koegl, Webster, & Levene, 2001). These widely used protocols assist clinicians in assessing the best known predictors of a child patient's prognosis, violence risk, and treatment engagement versus resistance, for case management purposes. However, diagnosing subtypes based on age of onset can be difficult during two developmental periods, childhood and adolescence, albeit for different reasons.

First, when a child presents for assessment, the task is to make a differential diagnosis between childhood-onset CD that will be only childhood-limited, versus childhood-onset CD that will follow a persistent course toward adult antisocial personality disorder. The age-of-onset distinction cannot help with this task because all child patients, by definition, have childhood onset. In the absence of advance knowledge of long-term persistence, the childhood-onset diagnosis does not reliably identify which CD children most need intervention, and may result in false-positive predictions of antisocial prognosis (Kim-Cohen et al., 2005; Lahey et al., 2005; Maughan & Rutter, 2001; Tremblay, 2000). Researchers have tried to distinguish life-course persistent versus childhood-limited trajectory groups by using childhood risk factors, without success (for a review see Moffitt, 2003, 2006).

Second, when an adolescent presents for assessment, the task is to make a differential diagnosis between adolescent-onset CD, versus childhood-onset CD that is already persistent and well on the way to a pathological adult prognosis. It seems obvious to ascertain age of first symptom to make the

subtype diagnosis, but this is easier said than done. A clinician may lack access to information about an adolescent patient's symptom history; credible reporters about the adolescents' childhood behavior may not be available, and even if reporters are at hand, retrospective reports are famously subject to memory failure (Simon & VonKorff, 1995). Research has shown that the age of onset of conduct problems is generally recalled as years later than it truly was (telescoping) (Henry et al., 1994). Official records of age at first police arrest also lag 2 to 5 years behind true age at first illegal act (Moffitt, Caspi, Rutter, & Silva, 2001). Good age-of-onset information is hard to get.

Possible disadvantages of onset-age subtypes

Although the adolescent-onset versus childhood-onset persistent CD subtypes offer good predictive validity and clinical utility, the current subtype system has disadvantages, as noted above. First, differential diagnosis of age-of-onset subtypes in adolescent clinical settings is not as reliable as it should be. Second, initial evidence suggests that the childhood-onset subtype may need to be further divided into life-course persistent versus childhood-limited groups, but it is not yet clear whether this division is necessary or how to accomplish it. Some but not all studies identify the childhood-limited group. Moreover, there is no consensus about whether the childhood-limited group's adult prognosis is good or poor, and if poor, what clinical outcomes the prognosis involves.

What research on the CD subtypes is needed?

First, research is needed to clarify the nature of the putative childhood-limited subtype of CD. Basic descriptive information is needed, as well as consensus about whether outcome always involves depression, anxiety, and social isolation. Short-term and long-term follow-up studies of both clinical and cohort samples are needed.

Second, given that age of onset is uninformative for subtyping child patients and tricky to assess in adolescent patients, research is needed to identify other factors that diagnosticians can use to differentiate between CD subtypes. These should be enduring characteristics of the child, not dependent on retrospective assessment. Possibilities include comorbid ADHD, biomarkers, psychopathic traits, male sex, and family history. Later sections of this paper will review research into these possibilities. Research is needed to test whether these factors differentiate CD subtypes in clinical and forensic settings.

Third, we need to ascertain how well results of trajectory analyses map onto DSM-IV CD. For example, do trajectory analyses support the age-10 cut off between the DSM-IV subtypes, or suggest a

new age cut-off? Also, how many members of trajectory groups would meet formal criteria for CD? In the Dunedin cohort, 100% of childhood-onset persistent trajectory members and 89% of adolescent-onset trajectory members met criteria for at least mild CD (3 + symptoms). More comparisons between trajectories and diagnoses are needed.

Fourth, we need information about how the subtypes might inform treatment choice, and predict treatment response. Much has been learned about the differential etiology of age-of-onset CD subtypes, and their differential long-term prognosis when left untreated, but very little is known about how these fundamental differences relate to intervention. Now that clinicians can choose among many different effective treatment approaches for CD (NIH State of the Science, 2004), any subtype system in DSM-V should help clinicians decide which treatment to choose. Future intervention research should compare the age-of-onset CD subtypes on key treatment measures.

Fifth, research is needed to determine whether or not developmental subtypes apply to children and adolescents from ethnic minority groups. Longitudinal cohort research has documented that the childhood-onset persistent and adolescent-onset subtypes apply to girls as well as boys (see this article's section on female-specific CD protocols), but findings with African-American cohorts are only suggestive (Moffitt, 2006), and other ethnic groups in the USA and other countries have not been studied.

Issue 2, family psychiatric history as an aid to diagnosing CD?

Definition of the issue

DSM-IV does not currently include family history information among the criteria for CD. However, family history information is routinely queried as part of the diagnostic protocol for many medical illnesses (Hunt, Gwinn, & Adams, 2003) and has been considered as a potential criterion for psychiatric illnesses such as depression (Kendler & Roy, 1995; Zimmerman et al., 2006). Conduct problems are known to be concentrated in families, and a family history of antisocial behavior is a robust predictor of offspring conduct problems (Farrington et al., 2001; Farrington & Loeber, 2000). Thus, the primary questions for this section are: Can family psychiatric history assist clinicians to predict CD prognosis or diagnose subtypes? If so, should family history be considered for inclusion in DSM-V?

Rationale for considering family psychiatric history for DSM-V

Three findings support the inclusion of family-history information in the CD criteria for DSM-V. First, family liability to antisocial behavior is at the

etiological core of CD. Meta-analysis of behavioral genetic studies has shown that CD is under moderate genetic influence (Rhee & Waldman, 2002). Genetic influence is very strong for the particular subtype of CD that has an early age of onset and is pervasive, persistent and severe (Arseneault et al., 2003; Moffitt, 2005a). Research on how genes may contribute to CD is well under way (Moffitt, 2005b) and genetic testing has been proposed for future classification systems (Charney et al., 2002). Taking a cautious view, however, it is unlikely that genetic markers will be included in DSM-V. In contrast, family history assessments are routinely used in medical practice to improve prediction of disease prognosis (Yoon, Scheuner, & Khoury, 2003) and are more feasible than genetic tests for marking the familial transmission of CD risk. Family history assessments can be powerful predictors because they have the advantage of comprising information about two causes of childhood-onset CD: familial genetic loading plus parents' environmental influences on their children's conduct.

Second, family history information may help clinicians make differential diagnosis between CD subtypes. Accurately diagnosing CD is complicated by the fact that the majority of children and adolescents who exhibit CD symptoms desist before they reach young adulthood and do not achieve the same level of poor prognosis as their childhood-onset life-course persistent counterparts (Robins, 1978; Rutter, Kim-Cohen, & Maughan, 2006; Tremblay et al., 2004). As noted in this article's section on CD subtypes, when a child presents for assessment, the clinician must make a differential diagnosis between childhood-onset CD that will be only childhood-limited, versus childhood-onset CD that will in future have a life-course persistent course and pathological prognosis. Likewise, when an adolescent presents for assessment, the clinician must make a differential diagnosis between adolescent-onset CD, versus childhood-onset CD that is already persistent and well on the way to a pathological adult prognosis. Longitudinal research has identified risk factors that reduce false positives in research settings (for a review see Moffitt, 2006), but many of these factors cannot be reliably or economically assessed in clinical settings. Initial research suggests that family history information may be able to resolve the clinician's subtyping dilemma (Odgers et al., in press). This research will be described below.

Third, knowledge of parental history of antisocial behavior may inform treatment planning for CD children. Parents with histories of CD provide suboptimal parenting and chaotic home environments (Jaffee et al., 2006). Children of antisocial mothers grow up in caregiving environments characterized by physical abuse, exposure to domestic violence, and maternal hostility (Kim-Cohen et al., 2006). The most effective interventions for CD invariably require the participation of parents (McCart et al., 2006), but

antisocial parents are at highest risk for terminating treatment (Kazdin, Holland, & Crowley, 1997). Thus, knowledge of family history could provide useful information to clinicians about both the child's and the family's potential amenability and responsivity to treatment.

Predictive validity of family psychiatric history

The Dunedin cohort study reported that a family history of externalizing disorders assessed in parents and grandparents (particularly conduct disorder/antisocial personality disorder; alcohol abuse; drug abuse) characterized the CD subgroup with a childhood-onset and subsequent persistent course of antisocial behavior to adulthood. However, family history did not characterize the CD subgroups which were childhood-limited or adolescent-onset. In this study, family history identified the childhood-onset persistent CD subtype that needed treatment most, with negative and positive predictive values of .70, and specificity of .95, suggesting little risk of false positives. Family history also provided incremental prediction for poor prognosis over and above CD symptom levels and key childhood-risk factors, including ADHD (Odgers et al., 2007). A similar finding was reported from a Minnesota sample; early-starter delinquents had more relatives who were offenders, as compared to late-starter delinquents (Taylor, Iacono, & McGue, 2000).

Reliability of assessing family psychiatric history

Brief family history assessments are routine in clinical settings when screening for medical diseases (Hunt et al., 2003) and have a relatively long history within psychiatric research settings (Andreasen et al., 1977). In research settings, family psychiatric health is commonly assessed via instruments such as the Family-History Screen (FHS), which has acceptable psychometric properties (Weissman et al., 2000). Moreover, clinicians working in pediatric treatment settings are often involved with multiple informants when treating a child (e.g., case workers, teachers, and caregivers) and, therefore, are likely to have access to information – albeit limited – regarding family members' history of antisocial behavior.

Possible disadvantages of a family psychiatric history criterion

Clinicians are unlikely to have time or resources to carry out detailed family history interviews, reliable informants may not be available (particularly in high-risk or juvenile justice settings), and social stigma may lead parents to under-report family history of criminal behaviors (Andreasen et al., 1977; Thompson et al., 1982). Thus, it may be difficult to obtain reliable family history information in clinical or forensic settings, as compared to in research. To

meet clinicians' needs, the above-mentioned Dunedin family history study tested the minimal subset of items and reporters that would yield useful diagnostic information about children's CD. Mothers' responses to three items about alcohol problems among the child's parents and grandparents discriminated between CD subgroups as well as the full family history interview, and helped to identify the childhood-onset life-course persistent CD subgroup that required treatment most (Odgers et al., 2007). Because alcohol problems are not inherently illegal, it may prove more practical to assess family histories of alcohol problems than family histories of criminal behavior.

What research on family psychiatric history is needed?

Family history information has demonstrated potential for improving prognosis prediction in research settings, but replication is needed to evaluate whether this initial finding extends beyond the Dunedin study. In preparation for DSM-V, epidemiological studies are needed to: (1) assess the sensitivity and specificity with which family history predicts CD, (2) estimate how many false-positive CD diagnoses would result from using family history, (3) determine which family members' psychiatric history should be assessed, and (4) evaluate whether a specific disorder among family members best predicts poor CD prognosis. Since data on family history are routinely collected in psychiatric epidemiologic studies, secondary analyses of existing data sets may address these issues. Research in clinical settings is also needed to determine (5) whether family psychiatric history helps to predict long-term prognosis and treatment response and, if so, (6) whether and how family psychiatric history can be assessed reliably by clinicians working in treatment settings.

Issue 3, callous-unemotional traits as criteria for CD?

Definition of the issue

The publication year of DSM-IV coincided with the first published research that attempted to extend the construct of psychopathy to children (Frick, O'Brien, Wootton, & McBurnett, 1994). The psychopathic personality in children was operationalized as 'callous-unemotional traits'. Regardless of whether callous-unemotional traits index later psychopathy, this line of research has evolved to identify a subgroup of children with CD who have a distinct neurological profile and worse prognosis than CD children without callous-unemotional traits. Some symptoms of these traits are mentioned in DSM-IV as an 'associated feature' of CD (p. 87). The question here is whether callous-unemotional traits should be added to DSM-V in some more formal way. This

addition could be accomplished in two ways. First, callous-unemotional features could be used as a criterion for subtyping CD children, much as psychopathy scores are sometimes used by researchers to subtype adults within the diagnosis of ASPD. Second, callous-unemotional trait behaviors could be added to the existing list of CD criterion symptoms. Both options should be evaluated.

Rationale for considering callous-unemotional traits for DSM-V

Current findings suggest that callous-unemotional traits may define CD children who have extreme behavior problems, stronger genetic risk, and at-risk neurocognitive profiles, when compared to other children with CD. Children with callous-unemotional traits show more conduct problems, more severe aggression and more proactive aggression than other children with CD (Frick & Marsee, 2006). Antisocial behavior is more heritable among children with callous-unemotional traits versus among other children (Viding, Blair, Moffitt, & Plomin, 2005). Children with callous-unemotional traits also show a specific neurocognitive profile suggestive of amygdala/orbitofrontal dysfunction, as manifested by insensitivity to punishment and distress cues (Blair, Peschardt, Budhani, Mitchell, & Pine, 2006; Dadds et al., 2006). The neurocognitive profile is similar to that seen in adult psychopaths (Lynam & Gudonis, 2005). Further, the profile differs from that of other children with CD, who do not show comparable punishment insensitivity, and, if anything, can be hypersensitive to anger and punishment cues (Blair et al., 2006).

Callous-unemotional traits may also be important when implementing treatment with CD children. In one study, children with callous-unemotional traits did not benefit from punishment-oriented behavior modification programs such as 'time-out,' which appear to work for other children with CD (Hawes & Dadds, 2005). There is also some evidence that children with CD and callous-unemotional traits may be less responsive to typical parental socialization practices than other children with conduct problems (Hipwell et al., 2007; Oxford, Cavell, & Hughes, 2003: Wootton, Frick, Shelton, & Silverthorn, 1997), possibly because they are less distressed by the effect their behavior has on others (Pardini, Lochman, & Frick, 2003).

Predictive validity of callous-unemotional traits

There is a relative lack of longitudinal follow-up data on the unique, incremental predictive value of callous-unemotional traits for antisocial outcomes (Frick & Dickens, 2006). However, the available evidence indicates that these traits index a relatively stable characteristic that predicts poor outcome. For example, callous-unemotional traits predict more

severe antisocial acts, delinquency, and higher rates of recidivism for adolescent offenders (Frick & Dickens, 2006; Frick, Stickle, Dandreaux, Farrell, & Kimonis, 2005; Forth, Kosson, & Hare, 2003). In one longitudinal study, callous-unemotional traits emerged alongside depression and drug use as the strongest predictors of later antisocial behavior (Burke, Loeber, & Lahey, 2007; Loeber, Burke, & Lahey, 2002). In another longitudinal study, psychopathy ratings at age 13 years predicted adult psychopathy status 11 years later, over and above prediction by other age-13 conduct problems (Lynam, Caspi, Moffitt, Loeber, & Stouthamer-Loeber, 2007).

Reliability of assessing callous-unemotional traits

The most widely used and best validated research measures of callous-unemotional traits are the Antisocial Process Screening Device (APSD; Frick & Hare, 2001), the Child Psychopathy Scale (CPS; Lynam, 1997), and the Psychopathy Checklist:Youth Version (PCL:YV; Forth, Kosson, & Hare, 2003). Different instruments have been used to assess callous-unemotional traits across studies, which has the advantage of showing the construct is robust, but may have the disadvantage of making comparison across samples difficult (Frick & Marsee, 2006; Lynam & Gudonis, 2005). We focus on the APSD and CPS here because they were designed to be used with young children.

The APSD callous-unemotional trait scale shows acceptable internal consistency (teacher $\alpha = .75$, parent $\alpha = .70$) (Frick & Hare, 2001). Inter-rater agreement (teacher-parent) is in the range of .30–.40, which is typical of most behavioral rating scales (Frick & Hare, 2001). Test–retest reliability coefficients are only available for teacher ratings of the APSD; and these range from .73–.87. Callous-unemotional traits as measured by the APSD showed considerable stability ($r > .70$ across all time points) over a four-year period in a small longitudinal study (Frick, Kimonis, Dandreaux, & Farrell, 2003).

The CPS total score shows excellent internal consistency (parent $\alpha = .91$). Individual subscales of the CPS are fairly reliable. However, the subset of items that specifically assess callous-unemotional traits shows more modest reliability. Parent–adolescent inter-rater agreement on the total score is .37, which again is typical of most behavioral scales. Test–retest reliability for the CPS has not been reported. A long-term follow-up study using the CPS at age 13 years demonstrated that the psychopathy construct showed moderate stability to age 24 years ($r = .32$), despite different informants and assessment instruments used across the two age periods (Lynam et al., 2007). This 11-year correlation is equivalent to that typically seen when different informants use the same instrument at the same time-point to rate an individual's behavior on a construct.

Currently, callous-unemotional traits are not routinely assessed when making a clinical diagnosis of CD. The assessment of callous-unemotional traits in research settings has generally relied on parent and teacher ratings. Such assessments could be easily adapted for clinical use to supplement standard diagnostic interviews.

Possible disadvantages of considering callous-unemotional traits in CD

Critics express concern about labeling children as having callous-unemotional traits or psychopathy. This concern emerges mainly because adult psychopaths are presumed to be untreatable. Clinicians rightly wish to avoid applying a label to children that implies they cannot be treated. However, identifying and treating callous-unemotional traits in children offers an important opportunity for prevention (see Lynam et al., 2007, for a discussion of this issue). Labeling is clearly a legitimate concern. However, this concern should not curb research that may lead to treatment options for vulnerable youngsters. Personality is more malleable in childhood than during older developmental stages (Roberts & DelVecchio, 2000). Thus treatment of psychopathic traits may be more effective for children than adults.

What research on callous-unemotional traits is needed?

First, psychometric evaluations in clinical settings are needed to assess the reliability of callous-unemotional assessment tools in these settings. Longitudinal follow-ups of clinical samples assessed for callous-unemotional traits should test whether they inform course and prognosis in clinical settings. Such research should clarify whether callous-unemotional traits have better predictive validity when used as a category or as a continuum. Because callous-unemotional traits are a relatively novel construct in child psychopathology research, continued psychometric work on optimizing the measurement of core callous-unemotional traits is warranted. In past versions of the DSM, observable behaviors were favored and personality traits were eschewed in the interest of attaining diagnoses with strong reliability; callous-unemotional traits were excluded from the criteria for CD and ASPD on that basis. Research will need to show that callous-unemotional traits improve the CD diagnosis, and do not reduce the internal consistency reliability of the CD construct.

Second, research is needed to ascertain any relation between callous-unemotional traits and the current subtypes of CD defined on the basis of age of onset. Extreme aggression, neurocognitive deficits, and poor prognosis characterize children with callous-unemotional traits, which suggests the

hypothesis that these traits typify the childhood-onset life-course persistent CD subtype. However, this hypothesis has not been directly tested. The question is important because CD already has a subtyping system, and we must know whether childhood-onset persistent children and callous-unemotional children are the same children. If callous-unemotional traits were deemed a useful addition to DSM-V, then research must determine if we should add callous-unemotional trait behaviors to the existing CD criteria, or alternately, consider callous-unemotional traits as an aid for subtyping.

Third, callous-unemotional traits should be incorporated into intervention research. There is a widespread belief that adult psychopaths are untreatable, making the notion of extending the psychopathy construct to children reprehensible to many mental health professionals. However, current treatments for CD may not meet the needs of children with callous-unemotional traits. Specifically, punishment-based approaches may not work optimally. Translational research is needed to develop and evaluate treatments incorporating strict boundaries, consistent rewards, and appeal to self-interest. Callous-unemotional traits should be studied as a moderating factor for response to current CD treatments.

Fourth, more research is needed into girls with callous-unemotional traits. Because most studies focus solely on boys, it is not clear whether callous-unemotional traits or psychopathy ratings capture the same latent constructs in boys and girls. There is some suggestion that girls with callous-unemotional traits engage in more relational vs. overt aggression (Penney & Moretti, 2006), and callous-unemotional relational aggression may be particularly related to girls' victimization experiences (Odgers, Reppucci, & Moretti, 2005). Research is also needed to ascertain whether callous-unemotional traits have good construct validity among ethnic minority children.

Fifth, research to identify indicators of callous-unemotional traits in preschool-aged children is needed. There is some suggestion that fearlessness in early childhood may be one index of risk for callous-unemotional traits and conduct problems (Frick & Marsee, 2006; Kochanska, Gross, Lin, & Nichols, 2002).

Sixth, epidemiological cohort studies are needed to report the prevalence of 'abnormal' callous-unemotional trait scores in the healthy, non-CD population. For example, teachers' and mothers' ratings of lack of remorse, 'does not seem guilty after misbehaving,' applied to 9–14% of 10-year-olds in the representative E-risk birth cohort, most of whom did not meet criteria for CD. If callous-unemotional traits were used routinely to aid diagnosis, what rate of false positives would be expected, and how could the rate be reduced?

Issue 4, neuroimaging biomarkers for CD?

Definition of the issue

The current DSM-IV diagnostic criteria are based on ratings of behavior. However, there is much excitement about discovering neuroimaging biomarkers relevant for psychiatric disorders. Neuroimaging biomarkers may potentially shed light on mechanisms connecting genes to behavioral disorders, including CD. Also, neuroimaging biomarkers have potential to differentiate subtypes within a heterogeneous disorder such as CD, and to reveal brain mechanisms uniting related disorders such as ODD, CD, and ASPD (Gould & Gottesman, 2006; Meyer-Lindenberg & Weinberger, 2006; Viding & Blakemore, 2006).

Only non-invasive neuroimaging methods can be used in children and these include: structural Magnetic Resonance Imaging (MRI), functional Magnetic Resonance Imaging (fMRI), Magnetoencephalography (MEG), Electroencephalography (EEG), and Event Related Potentials (ERP). Of these, all but MEG have been used to study CD/antisocial behavior in childhood. The question is, should neuroimaging biomarkers be recommended as a complementary assessment method for CD?

Rationale for considering neuroimaging biomarkers for DSM-V

The existing, small neuroimaging literature is suggestive of frontal and temporal abnormalities in CD. ERP studies show that children with CD have reduced P300 amplitude in executive and monitoring tasks in anterior brain sites (e.g., above the anterior cingulate cortex) (Bauer & Hesselbrock, 2003; Kim, Kim, & Kwon, 2001; Costa et al., 2000). A longitudinal study using ERP measures found that future criminal offenders were characterized by larger N1 amplitudes and faster P300 latencies to the warning stimulus in a reaction time task, compared to controls (Raine, Venables, & Williams, 1990). Data from fMRI studies show that CD children have reduced anterior cingulate responsivity to emotional stimuli, which is taken to reflect poor emotional regulation (Sterzer et al., 2005; Stadler et al., 2006). In addition, CD children show amygdala hyporeactivity to emotional stimuli when variance associated with anxiety is partialed out of the analysis (Sterzer et al., 2005). MRI studies have identified structural abnormalities in children with CD, including abnormalities in temporal gray matter volume (Kruesi et al., 2004) and white matter hyperintensities in frontal lobes (Lyoo et al., 2002). Finally, the EEG data suggest that aberrant resting activity in left frontal sites is associated with retrospectively rated childhood CD symptoms (Deckel, Hesselbrock, & Bauer, 1996). These initial findings of structural and functional abnormalities in frontal and temporal regions are consistent with the neuropsychological

literature showing that children with CD have deficits in executive function and affective processing (e.g., Nigg & Huang-Pollock, 2003; Blair et al., 2006).

Predictive validity of neuroimaging biomarkers

Only one study to date has looked at the predictive validity of neuroimaging biomarkers in CD. Raine et al. (1990) collected ERP measurements from 101 boys at age 15 and assessed their criminal behavior at age 24. Criminals-to-be were characterized by larger N1 amplitudes and faster P300 latencies to the warning stimulus in a task where they had to be ready to react to an impending stimulus. Clearly, more work is needed to evaluate the predictive validity of neuroimaging for CD children's outcomes.

Reliability of assessing neuroimaging biomarkers

Although test–retest reliability data on neuroimaging measures exist (e.g., P300 amplitude and latency show high test–retest reliability above .80; Hall et al., 2006), there is no reliability data on clinical assessment because neuroimaging measures are not yet used to complement symptom diagnosis. The test–retest reliability for MRI structural data and fMRI paradigms runs from adequate to excellent (most reported values in the range of $r = .60–.90$), including reasonable test–retest correlations for left amygdala activity (.63–.70; Johnstone et al., 2005). Although calibrating signals across machines at different research sites can be difficult, data are emerging to show that different laboratories running similar fMRI paradigms usually report comparable sites of brain activation (e.g., Phan, Wager, Taylor, & Liberzon, 2002).

Possible disadvantages of using neuroimaging biomarkers

The current neuroimaging literature related to CD is still sparse, although it is growing. Most neuroimaging samples have been small and subject to selection biases. These small samples are further hampered by heterogeneity because studies have not typically differentiated between CD subtypes. In addition, it is unclear how specific the reviewed neuroimaging biomarkers are to CD. Frontal lobes cover one-third of the brain and several other childhood disorders are associated with frontal lobe dysfunction. Likewise, CD is not the only disorder associated with temporal lobe abnormalities, even if we narrow the abnormality to include the amygdala only.

Neuroimaging methods provide potentially useful information about the mechanisms of a disorder and have strong promise for resolving heterogeneity into subtypes. However, from a clinical point of view, most current neuroimaging methods are not feasible in clinical practice. These measures require intensive specialized training, are very expensive, and are

not currently reimbursed by insurance as a diagnostic procedure. Other considerations include the suitability of fMRI with girls who may be pregnant or with children who have been shot or stabbed and might have metal in their bodies.

What research on neuroimaging biomarkers is needed?

Replications of the initial findings of neuroimaging biomarkers for CD are needed, and effect sizes for imaging measures should be estimated through meta-analyses. The predictive validity of neuroimaging biomarkers needs to be tested in longitudinal follow-ups of at-risk children who have taken part in imaging studies. The imaging paradigms will also need to be standardized across laboratories to meaningfully interpret and replicate findings related to CD. In addition, neuroimaging biomarkers must be assessed in epidemiological cohort samples in order to address unanswered questions about their sensitivity and specificity to CD (or a particular subtype of CD). For example, it is important to assess potential effects of age, sex, and race differences on these markers, and to estimate the prevalence of 'abnormal' biomarker scores in healthy, non-CD children.

Translational research is needed to develop more accessible and easy-to-employ or automated neuroimaging methods that would be available for wider clinical use. Once protocols are developed, their psychometric properties, such as reliability, should be tested.

Given that children with CD are a heterogeneous group, researchers who study neuroimaging biomarkers should pay attention to CD subtype distinctions in order to prevent sample heterogeneity from masking true associations between neuroimaging biomarkers and specific CD subtypes. As an example, existing neuropsychological research in children suggests that neuroimaging biomarkers may be quite different for CD children with anxiety vs. CD children with callous-unemotional traits; the former show hypersensitivity to threat stimuli and the latter show hyporeactivity to others' distress (Blair et al., 2006). In addition to CD subtype distinctions, researchers should pay attention to gender differences. Currently it is unclear how well the imaging research applies to girls with CD. The majority of the participants in imaging studies have been males.

Ultimately, neuroimaging research may be helpful in validating subtypes, or in guiding treatment development. The strength of neuroimaging approaches is that they enable researchers to access preconscious, automatic emotional and attentional processes and relate these to behavior. As such, neuroimaging research could eventually help clinicians to design treatment approaches that circumvent a patient's impaired brain capacities and draw

on the patient's unimpaired brain capacities. Neuroimaging can also be used to assess functional brain changes in response to treatment, an approach that has not yet been applied to the study of CD. For now, we suggest the emphasis should be placed on using neuroimaging biomarkers to better understand the causal mechanisms involved in CD, rather than to become part of routine diagnostic assessment.

Issue 5, genotypes as biomarkers for CD?

Definition of the issue

Research on how genes may contribute to CD is well under way (Moffitt, 2005a, 2005b) and the inclusion of genetic testing has been proposed for future classification systems (Charney et al., 2002). As knowledge builds regarding gene-to-behavior associations, it becomes appropriate to ask whether reliable and valid genotypic indicators of CD are available for improving diagnostic accuracy in DSM-V.

Rationale for considering genotypic biomarkers of CD for DSM-V

Antisocial disorders are concentrated in families (see section on family history, this article), and there is now solid evidence from twin and adoption studies that conduct problems are at least moderately heritable (Moffitt, 2005a; Rhee & Waldman, 2002).

Studies of gene-to-CD associations consist of two types. First, 'main effect' studies investigate direct associations between genetic variants and CD using candidate-gene association approaches. Second, genotype-by-environment interaction (G × E) studies investigate genetic moderation of environmental effects on CD. Suggestive evidence for both types of associations has emerged. Several recent candidate gene studies have reported a main effect association between polymorphisms in the serotonin transporter gene *5HTTLPR* and conduct problems (Beitchman et al., 2006; Haberstick, Smolen, & Hewitt, 2006; Sakai et al., 2007). Genes in the dopamine neurotransmitter system that have been implicated in conduct problems include the dopamine receptor *DRD4* (Holmes et al., 2002), the dopamine transporter *DAT1* (Young et al., 2006), and the catechol *O*-methyltransferase gene *COMT* (Caspi et al., in press; Thapar et al., 2005). However, not all replications of these findings are positive (e.g., Sengupta et al., 2006).

A GxE study showed that a polymorphism in the *MAOA* gene significantly moderates the impact of childhood maltreatment on risk for CD, aggression, and violent crime in adolescence and adulthood (Caspi et al., 2002). The gene encodes the *MAOA* enzyme which selectively metabolizes serotonin, norepinephrine, and dopamine – neurotransmitters important for the regulation of mood and behavior. An initial meta-analysis of five studies showed that the interaction between *MAOA* genotype and childhood maltreatment predicting conduct problems was modest but statistically significant (Kim-Cohen et al., 2006). Pooling the five samples, the correlation between child maltreatment and conduct problems was .12 in individuals with the high-activity *MAOA* genotype, and .32 in individuals with the low-activity genotype. New studies of this G × E hypothesis are appearing (a negative replication by Huizinga et al., 2006; a positive replication by Widom & Brzustowicz, 2006). With these two studies added to the aforementioned meta-analysis, the correlation between maltreatment and conduct problems was .13 in high-activity genotype individuals, and .30 in low-activity genotype individuals (Taylor & Kim-Cohen, 2007). Meta-analyses should be updated to accommodate new data about genetic biomarkers as they emerge.

To date, three genome-wide linkage studies of CD have been conducted (Dick et al., 2004; Kendler et al., 2006; Stallings et al., 2005). Each tentatively suggests regions on chromosomes that might harbor CD-related genes (specific genes must now be identified in these regions). However, the three studies have not converged on the same chromosomal regions, except possibly regions 1q and 2p. These genome-wide scans were carried out in samples originally ascertained to study people at high risk for substance dependence. Such samples may not represent CD that occurs in the absence of familial substance dependence. However, as noted in this article's section on family history, a familial liability to substance abuse particularly characterizes the childhood-onset life-course persistent subtype of CD, and thus the combined CD-substance-risk phenotype may be ideal for gene-finding.

Predictive validity of genotypic biomarkers of CD

Because genotypes do not change over time, all studies of the above-mentioned genotype-to-behavior associations can be considered to be tests of predictive validity (i.e., temporal order between DNA and behavior is never in question). Beyond this, evidence that genetic markers predict CD with specificity is lacking. For instance, *MAOA, DRD4*, and *DAT1* have been associated with ADHD (Thapar, O'Donovan & Owen, 2005), *5HTTLPR* has been associated with depression (Hariri & Holmes, 2006), and *COMT* has been associated with psychosis (Craddock, Owen, & O'Donovan, 2006). Moreover, no studies have been conducted to show whether genetic markers of CD can improve prediction of prognosis above and beyond conventional symptom information. Finally, no studies have tested if genetic markers can identify CD individuals who need treatment most or who might benefit from types of

treatment, including personalized prescribing of pharmacological treatments.

Reliability of genotypic biomarkers of CD

Although genotyping error occurs (Pompanon et al., 2005), genotyping is no less reliable than other medical laboratory tests now in routine use. However, it is not yet clear how genotyping or genetic assessments can be incorporated into clinical assessment contexts. Unlike neuroimaging, genotyping need not be costly, but clinicians in most settings do not have access to genotyping facilities nor do insurance carriers reimburse for genotyping at this time. In addition, new genetic-counseling protocols would be needed before clinicians could ethically communicate to families about probabilistic genetic risk for complex childhood disorders such as CD. Therefore, it is not currently feasible for genetic markers to be used to aid diagnosis in typical clinical settings.

Possible disadvantages of genotypic biomarkers of CD

The current evidence base reveals many disadvantages in the use of genotypic biomarkers for diagnostic purposes. Application is premature because the replication record is still under construction, and the strength of association between individual genes and psychiatric disorders is weak, often nonspecific, and embedded in complex causal pathways involving other, non-genetic influences (Kendler, 2005). Susceptibility genes, even if they were to be reliably identified, can vary across individuals in the probability that the disease phenotype will be expressed (Merikangas, 2002). G × E findings suggest this varying connection to behavior may depend in part on individuals' exposure to the environmental risk factors for CD. Moreover, the frequency of genetic variants and their functions may differ according to age, ancestry and sex (Charney et al., 2002). Therefore, conclusions regarding genotypic markers for CD may not apply to some population subgroups, but specifics remain unknown at present. Finally, labeling at-risk individuals on the basis of their genotype can be stigmatizing, and legal protections against potential discrimination and misuse of genetic diagnostics are not in place.

What research on genotypic biomarkers of CD is needed?

The search for specific genetic polymorphisms associated with conduct problems is still a very new initiative and the existing evidence base does not yet support the use of genetic markers for the purpose of diagnosing CD. Taking a cautious view, it is unlikely that genetic markers will be included in DSM-V, because much more research is needed.

Well-designed whole-genome scans are needed to screen for potential new genetic variants associated with CD and its intermediate phenotypes. As new variants are identified, hypothesis-driven studies are needed to elucidate the biological mediators of links between candidate genes and CD. As one example, individuals who differ on the MAOA polymorphism linked to CD also differ on structural and functional MRI measures involving the amygdala and anterior cingulate (Meyer-Lindenberg et al., 2006). Because the environmental causes of CD are well known, both whole-genome scans and hypothesis-driven candidate gene studies should study samples exposed to environmental risk factors for CD. Such samples will help to determine which genetic variants characterize children who do versus do not develop CD when children are exposed to an environmental pathogen such as maltreatment (Moffitt, Caspi, & Rutter, 2005).

In general, non-replication of findings has hampered progress in psychiatric genetics. Thus, independent efforts to replicate published results are needed, and both positive and negative replication studies must be published in order to gain a complete understanding of the evidence base. As findings emerge, meta-analysis to estimate pooled effects across studies can bolster conclusions regarding genetic biomarkers for CD (Ioannidis et al., 2001).

Epidemiological studies will be needed to ascertain basic descriptive information about genetic biomarkers associated with CD before we can evaluate their potential clinical utility. Such questions include: Does the genetic marker show specificity to CD as compared to other psychiatric disorders? Are there potential differences in gene–CD associations across age, sex, race/ethnicity, and CD subtype groups? What is the prevalence of risk genotypes in the healthy, non-CD population? If a genetic test were ever used to aid diagnosis, what rate of false positives would be expected? Can taking a family history provide comparable information to genetic testing, without the inherent disadvantages?

Research is needed that asks whether identified genetic markers can improve prediction of prognosis. Can genotype help identify which subtypes of CD children need treatment most, or which CD patients might benefit from particular types of treatment?

Issue 6, physiological biomarkers for CD?

Definition of the issue

The DSM-IV diagnoses CD on the basis of observable behaviors, but a growing body of research indicates that physiological biomarkers are associated with CD. DSM-IV mentions only heart rate and skin conductance as 'associated laboratory findings' for CD (p. 88). This section reviews the physiological biomarkers that are most robustly associated with conduct problems: heart rate, stress hormones,

neurotransmitters, and perinatal complications. The question is, are any of these biomarkers ready to inform DSM-V?

Slow resting heart rate (or pulse rate) is the most replicated of all biological markers for conduct problems. A meta-analysis of 40 studies concluded that slow heart rate, both resting and during a stressor, is a robust correlate of conduct problems (Ortiz & Raine, 2004). Effect sizes were moderate for resting heart rate (–.44) and large for heart rate during a stress challenge (–.76). The activity of the hypothalamic-pituitary-adrenal axis has also been associated with conduct problems. Studies have shown that low resting and hyper-reactive levels of the stress hormone cortisol characterize children, adolescents, and adults with conduct problems (Susman, 2006). Effect sizes are moderate for both resting cortisol levels (–.40) and reactive cortisol level following a challenge (.42) (van Goozen, Fairchild, Snoek, & Harold, 2007). Other hormones such as testosterone have been linked to aggression and violent behavior (Nelson, 2006). Neurotransmitters have also been linked to conduct problems, particularly serotonin (van Goozen et al., 2007). A meta-analysis of 20 studies of adults and five studies of children reported a moderate association between reduced serotonin metabolite levels and conduct problems (Moore et al., 2002). Perinatal and obstetric complications have been associated with conduct problems across the life span (Brennan, Grekin & Mednick, 2003; Raine et al., 1994). Such complications are assumed to engender or signal fetal brain damage, which in turn may cause neuropsychological deficits that are known risk factors for conduct problems.

Rationale for considering physiological biomarkers of CD for DSM-V

Two reasons have been suggested for considering physiological markers as potential diagnostic criteria for DSM-V CD (Popma & Raine, 2006). First, physiological markers (like genes and neuroimaging) may confirm homogeneous subtypes of CD individuals with characteristic pathophysiological profiles. Disorders may also be grouped in DSM-V on the basis of shared biomarkers; groupings are currently defined on the basis of surface symptom similarity alone, and the resulting errors retard research (Phillips et al., 2005). Second, physiological biomarkers (like genes and neuroimaging) could bring objective and unbiased information to the diagnostic process, reducing diagnosticians' exclusive reliance on parents' and children's reports.

Predictive validity of physiological biomarkers

Most studies on physiological markers are cross-sectional. Therefore, it remains unclear whether these markers can predict conduct problems or whether they are the consequence of an adverse, antisocial lifestyle. Only a few markers have been studied in the context of prospective longitudinal studies designed to test whether physiological markers predict later-emerging conduct problems. The exception is perinatal complications; because they occur at the very beginning of life, all relevant studies have tested whether perinatal measures prospectively predict later conduct problems.

Increasing evidence suggests that physiological markers operate in interaction with the social environment in predicting conduct problems (Raine, 2002). For example, studies have indicated that perinatal complications (e.g., minor physical anomalies, prenatal exposure to nicotine, obstetric delivery complications) predict conduct problems specifically in children and adolescents who were rejected by their mothers, who grew up in unstable families, or who were raised in deprived environments (Arseneault et al., 2002; Piquero & Tibbetts, 1999; Raine et al., 1994). Similarly, slow heart rate predicts later conduct problems over and above other risk factors, but findings further show that the highest probability of violence is among individuals who show a combination of slow resting heart rate and other social risk factors (e.g., large family size, poor relationship with parents) (Farrington, 1997).

Reliability of assessing physiological biomarkers of CD

For many physiological protocols, test–retest reliability has not been established. Moreover, measures and assessment methods are not consistent across studies because to date research teams follow different, tailored protocols. Few of the biomarkers mentioned above can easily be assessed by clinicians. Heart rate can be quickly and reliably assessed using non-sophisticated apparatus, although there is no established clinical cut-off for what constitutes 'slow' heart rate (Ortiz & Raine, 2004). However, the assessment of most other physiological biomarkers in the context of clinical settings and for clinical purposes is more complicated. Cortisol levels are also sensitive to circadian cycles, stress, diet, and physical exercise, they require laboratory expertise, and they are expensive to assay. Assessment of serotonin metabolites necessitates lumbar puncture, which may be painful and potentially harmful for children. Information about obstetric complications relies on hospital records but the accessibility, completeness, and reliability of those records varies across institutions.

Possible disadvantages of including physiological biomarkers in diagnosing CD

The feasibility of assessing most physiological markers in clinical settings is a major obstacle. First, there is no consensus as to when biomarker levels

are indicative of clinically significant diagnosed CD. Second, physiological biomarkers are sensitive to the influence of other factors such as child-rearing history, stress, diet, and exercise. Thus, including biomarkers among the diagnostic criteria for CD would mean assessing a wide range of potential confounding variables to aid interpretation of biomarker data. Third, some markers may have different correlates for males and females (e.g., testosterone) and little is known about age-related change or stability in biomarker levels.

What research on physiological biomarkers is needed?

Although assessing physiological markers may improve our understanding about the etiology and course of CD, a great deal of further research is needed before they can inform DSM-V diagnosis. First, the field is far from consensus about the role of these biomarkers in conduct problems. It is essential to develop standardized measures and assessment methods for research into biomarkers, to reconcile divergent findings across studies. Second, epidemiological studies are needed to examine the potential effects of age, sex, and race on the associations between physiological markers and CD. Third, psychiatric control comparisons are needed to determine whether physiological biomarkers show specificity to CD (or group it with etiologically related disorders). For example, low heart rate seems to be a specific risk factor for CD (Ortiz & Raine, 2004), but obstetric complications are not (Cannon et al., 2002). Fourth, longitudinal, prospective, and experimental research is needed to determine whether physiological biomarkers are true etiological factors for CD, or mere correlates, and how well they predict prognosis. Fifth, there is some evidence to suggest that biomarkers such as slow heart rate and perinatal complications differentially characterize the childhood-onset persistent subtype of CD, but more research is needed to ascertain whether other biomarkers relate to CD subtypes. Sixth, epidemiological studies are required to reveal the prevalence rate of abnormal physiological biomarkers in the healthy, non-CD population of children. Prevalence data will help anticipate the risk of false positives if biomarkers were used in diagnosis. Finally, if all of these hurdles were cleared by a biomarker associated with CD, translational research would be needed to convert research paradigms to protocols that are reliable, valid, and feasible for clinical practice.

Issue 7, diagnosing CD in preschool children?

Definition of the issue

The DSM-IV specifies that CD may onset in children as young as age 5–6 years but usually onsets later

(p. 89); DSM-IV makes no mention of whether CD may occur in preschool-aged children. Whether CD can be reliably and validly diagnosed in children ages 2–5 years has been a focus of controversy (Keenan & Wakschlag, 2002). The question for this section is, can and should DSM-V indicate the diagnosis of CD in preschool children?

Rationale for considering preschool CD diagnosis for DSM-V

Proponents of early-childhood diagnosis note that CD symptoms first emerge at preschool ages, timely intervention is desirable for both preschoolers and parents, and effective treatments are available. However, a barrier to affordable intervention for high-risk families is access to a diagnosis, which is conventionally required to secure clinical services.

The broader category of disruptive behavior disorders, which includes CD, is the top reason for referral of young children to mental health clinics (Keenan & Wakschlag, 2000). Population studies of preschoolers estimate the prevalence of DSM-IV CD from 3% to 7% (Egger & Angold, 2006; Kim-Cohen et al., 2005). Longitudinal follow-up studies of preschool cohorts have shown that 5–15% of 2–4-year-olds with high levels of aggressive conduct problems develop a persistent course and have poor prognosis many years later (NICHD Early Child Care Research Network, 2004; Campbell et al., 2006; Coté, Vaillancourt, LeBlanc, Nagin, & Tremblay, 2006). Preschool intervention is desirable to prevent chronic CD (Tremblay et al., 2004), and a diagnostic system is needed to determine a reliable and valid threshold for identifying which preschool-aged children require treatment.

Readers may wonder why a diagnosis of ODD is not sufficient to identify treatment needs at preschool ages. In preschool research samples, there is considerable overlap between CD and ODD groups. The ODD diagnosis captures most children who also meet CD criteria, but also quite a few other children who do not (Keenan et al., 2007). Preschool ODD has evidence of predictive validity for short-term prognosis (Speltz et al., 1999), but the CD group within ODD appears to be the subset most urgently in need of services (Keenan & Wakschlag, 2000). Very few studies of preschoolers have examined both ODD and CD in the same samples, and no conclusion can yet be reached regarding whether ODD is sufficient to identify treatment need at this age, or whether a CD diagnosis would improve clinical practice and service delivery (Egger & Angold, 2006).

Predictive validity of diagnosing CD in preschoolers

To date, only one epidemiological study has investigated DSM-IV diagnosed CD in non-referred preschoolers. This cohort study found evidence supporting the predictive validity of the prospective

DSM-IV CD diagnosis at age $4\frac{1}{2}$–5 years by following the cohort up to ages 7 and 10 years. Compared with non-diagnosed children, 5-year-olds diagnosed with CD were at significantly greater risk for a CD diagnosis when followed up at age 7 (OR = 20.6; 95% CI = 12.5–34.1). Conduct disordered 5-year-olds were also significantly more likely than non-diagnosed counterparts to have behavioral and educational difficulties at age 7. Although many 5-year-olds showed apparent remission from CD by age 7, these children continued to experience clinically significant behavioral and academic difficulties (Kim-Cohen et al., 2005). Diagnosed 5-year-olds' outcomes at age-10 follow-up were similarly poor as at age 7 (Kim-Cohen et al., in revision).

Reliability of preschool CD diagnosis

The evidence indicates that CD can be reliably diagnosed in the preschool period. The Kiddie-Disruptive Behavior Disorder Schedule (K-DBDS; Keenan et al., 2007) is a semi-structured parent interview that covers DSM-IV CD, ODD, and attention deficit hyperactivity disorder. It is based on the Kiddie-Schedule for Affective Disorders (K-SADS; Orvaschel & Puig-Antich, 1995), and modifications were made to provide developmentally appropriate operational definitions and to eliminate symptoms that lacked face validity (Keenan & Wakschlag, 2004). One-week test–retest reliability for the CD diagnosis is good (kappa = .73), and inter-rater reliability is excellent ($r = .96$) (Keenan et al., 2007).

The Preschool Age Psychiatric Assessment (PAPA; Egger et al., 2006) is a structured parent interview for diagnosing DSM-IV disorders, including CD, in children ages 2–5 years. The PAPA criteria for CD include 10 of the 15 DSM-IV-TR symptoms (5 symptoms were excluded as developmentally inappropriate: 'stealing with confrontation,' 'forced sexual activity,' 'breaking into a house or a car,' 'running away from home,' and 'truancy'). The presence of 3 or more symptoms qualifies for a diagnosis. One-week test–retest reliability is good (kappa = .60; intra-class correlation = .66), and comparable to that reported for structured interviews used to diagnose older children and adolescents (Egger et al., 2006).

Possible disadvantages of preschool CD diagnosis

The rationale against diagnosing CD in preschoolers comprises several different objections. Aggressive behavior is common and developmentally normative in the preschool period (Tremblay et al., 2004). Most preschoolers will naturally learn alternatives to aggressive behavior as they develop (Campbell, 2002), and the general population trend is for conduct problems to decrease across the first 10 years of life (Hill et al., 2006). The predictive

accuracy of conduct problems for future CD is thought to improve only when children grow older (Bennett & Offord, 2001). 'Down-aging' diagnostic criteria validated for older children and adolescents to young children may promote over-diagnosis and unnecessary treatment (McClellan & Speltz, 2003). Young children may be especially vulnerable to being labeled with a psychiatric disorder that may have adverse consequences for their self-perception and the perceptions of parents and other adults (Egger & Angold, 2006). Finally, diagnosing the problem in the child may overlook the fact that in very young children, the problem often lies in the relationships between children and parents.

What research on preschool CD diagnosis is needed?

First, research is needed to test the predictive validity of the CD diagnosis in preschoolers. Specifically, research using standardized diagnostic methods and representative samples of children aged 2–4 years is needed. Preferably, follow-up studies will track the outcomes of CD-diagnosed preschoolers over several years, into adolescence.

Second, more research is needed to gain consensus on whether and how diagnostic criteria for CD in preschoolers should be modified to be developmentally appropriate. Diagnostic protocols for CD in preschoolers include the PAPA and K-DBDS. Also, a set of developmentally appropriate diagnostic criteria has been recommended, called the Research Diagnostic Criteria-Preschool Age (RCD-PA; Task Force on Research Diagnostic Criteria: Infancy and Preschool, 2003). However, each of these systems proposes a somewhat different set of modifications to the DSM-IV criteria. The reliability and validity of both symptom definitions and duration criteria should be evaluated (Wakschlag, Leventhal, & Thomas, in press).

Third, children are generally thought to be unable to accurately report about their own behavioral and emotional symptoms. Hence, diagnostic tools for preschoolers have not utilized self-reports. Recent evidence, however, has indicated that using a developmentally appropriate instrument called The Berkeley Puppet Interview can yield reliable assessments of children's disruptive behavior (Arseneault et al., 2005; Measelle, Ablow, Cowen, & Cowen, 1998). The Berkeley Puppet Interview CD scale has test–retest reliability ranging from .52 to .69 (Ablow et al., 1999) and internal consistency reliability of .81 (Arseneault et al., 2005). The utility of incorporating preschoolers' self-reports into CD diagnosis should be explored.

Fourth, longitudinal epidemiological cohort studies are needed that test whether the diagnosis of CD in preschoolers adds incremental predictive validity over and above the diagnosis of ODD.

Issue 8, female-specific diagnostic protocols for CD?

Definition of the issue

The sex ratio for CD is approximately 2.5 males for each female (Moffitt et al., 2001). There is debate about whether the lower prevalence rate of CD among females reflects true sex-differences in CD versus sex bias against girls in the diagnostic criteria. DSM-IV mentions CD symptoms such as running away and prostitution as typical of girls, and notes that girls tend to use nonconfrontational aggression (p. 88). However, DSM-IV does not currently include sex-specific criteria for CD and the criteria were developed and validated primarily on male samples. Clinicians have become increasingly concerned about treating CD among girls and girls' CD is currently a topic of intense research (Moretti, Odgers, & Jackson, 2004; Pepler et al., 2005; Putallaz & Bierman, 2004). The question for this section is: Should DSM-V incorporate sex-specific aspects into the diagnosis of CD?

Rationale for considering female-specific diagnostic protocols for DSM-V

There are 3 main positions regarding the need for sex-specific diagnostic protocols for CD in DSM-V. First, girls with subclinical symptoms go on to experience clinically significant long-term outcomes and it may be necessary to lower the symptom threshold for females to ensure that these cases are not missed (Zoccolillo, 1993; Zoccolillo, Tremblay, & Vitaro, 1996). Second, the vast majority of female CD cases onset in adolescence, leading some researchers to argue that a modified version of the adolescent-onset subtype is sufficient to characterize girls' conduct problems, implying that a childhood-onset CD subtype is not needed for girls (Silverthorn & Frick, 1999). Third, the DSM-IV symptom criteria focus on behaviors that are more common among boys, and consequently the current CD criteria may fail to accurately detect CD among girls. Behaviors such as relational aggression, that are more common among girls than physical aggression, are not included in the CD criteria and researchers have argued for their inclusion (Crick & Zahn-Waxler, 2003). Girls' relational aggression shares correlates with boys' childhood-onset conduct problems (Marsee, Silverthorn, & Frick, 2005; Zalecki & Hinshaw, 2004), suggesting that, in the absence of overt DSM-IV CD symptoms, relational aggression may serve as an alternative marker of CD risk for girls (Frick & Dickens, 2006).

Predictive validity of female-specific diagnostic protocols

Would sex-specific protocols improve prediction of girls' prognosis and identify girls who need treatment?

First, some diagnosticians propose that DSM-V should lower the number of symptoms required to diagnose CD among girls. However, epidemiologic data demonstrate that subclinical conduct problems predict poor prognosis in adulthood among both males and females (Messer et al., 2006; Moffitt et al., 2001). To date, there is no compelling evidence for lowering the CD threshold for one sex versus the other. Rather, as detailed in this article's section on categorical versus continuum approaches, there is little evidence to support a categorical turning point along the distribution of CD symptoms for either males or females – suggesting that a dimensional operationalization of CD may be the best approach for both sexes. In any case, the issue of thresholds may assume less importance in the future as a result of the trend toward incorporating the dimensional approach to disorders in DSM-V.

Second, some diagnosticians propose that DSM-V should specify different subtypes of CD for girls and, specifically, that DSM-V may not need the childhood-onset type for females. However, a number of prospective cohort studies have identified a childhood-onset CD subtype of females who go on to experience poor prognosis in adolescence and adulthood (Coté, Zoccolillo, Tremblay, Nagin, & Vitaro, 2001; Fergusson & Horwood, 2002; Lahey et al., 2006; Odgers et al., in press; Schaeffer et al., 2006). Research also suggests that the origins of childhood-onset conduct problems are the same for the sexes, but that childhood-onset CD is more common among boys than girls because individual-level risk factors for childhood-onset CD (e.g., hyperactivity, verbal deficits) are more prevalent among boys than girls (Gorman-Smith & Loeber, 2005; Lahey et al., 2006; Messer et al., 2006; Moffitt et al., 2001).

Third, some diagnosticians propose that DSM-V should add female-specific symptom criteria. However, there are currently no empirical data available to evaluate whether the inclusion of sex-specific symptoms improves the prediction of girls' prognosis. Many of the symptoms suggested for girls, such as relational aggression, are highly correlated with other forms of aggression (Archer & Coyne, 2005; Odgers & Moretti, 2002; Xie, Swift, Cairns, & Cairns, 2002) and this redundancy implies that they may not provide incremental predictive validity for prognosis. In addition, some features of relational aggression are already in the DSM system under ODD, which includes criteria such as 'spiteful and vindictive.' Substance misuse has been suggested as particularly prognostic for girls, but this hypothesis needs more research (see this article's section on the early substance use as a criterion for CD).

Reliability of female-specific diagnostic protocols

Are female-specific criteria assessed reliably by clinicians? Evidence from clinical settings suggests

that sex-specific criteria are already being used to assess CD. For example, the *Early Assessment Risk List for Girls* (Levene et al., 2001) is a risk-assessment tool for girls exhibiting symptoms of CD prior to the age of 12. The EARL-21G was constructed by modifying item descriptions and adding two 'gender-sensitive' items ('Care-giver Daughter Interaction' and 'Precocious Sexual Development and Behavior') to the more widely used *Early Assessment Risk List for Boys* (Augimeri et al., 2001). To date, there is little empirical data available to evaluate the psychometric properties and predictive validity of the EARL-21G and, in general, the clinical utility of similar sex-specific initiatives remains an important, but still unanswered, question (Hipwell & Loeber, 2006; Levene et al., 2004).

Possible disadvantages of female-specific diagnostic protocols

The inclusion in DSM-V of sex-specific diagnostic protocols for CD will necessarily increase the prevalence rate of CD among girls. Lowering the cut-off threshold for diagnosis, or adding more female-specific symptom criteria, would increase the prevalence rate of CD but these changes might also improve detection of girls who need treatment. However, more research on the sensitivity and specificity of any sex-specific protocol is needed to ensure that diagnostic procedures do not over-identify girls who have little risk for poor prognosis.

What research on female-specific CD protocols is needed?

We noted three proposed sex-specific protocols for CD: lowering the diagnostic threshold, eliminating the childhood-onset subtype for girls, and adding female-specific criteria. Longitudinal cohort studies that have examined CD thresholds and developmental subtypes have not provided compelling support for these two sex-specific changes. In contrast, the proposal to add female-specific criterion symptoms to DSM-V remains under-researched.

Epidemiological research is needed to test whether proposed female-specific symptoms (such as relational aggression, substance misuse, risky sexual behavior, and conflict with caregivers) have specificity to CD. The female-specific symptoms should also transcend race and age, and they should not be so prevalent in the healthy non-CD population of girls that they increase the risk of false-positive diagnosis. In addition, research should check whether symptoms thought to be female-specific might be part of the CD construct for boys, as well as for girls. This research should comprehensively map the 'construct space' for CD by including DSM-IV CD symptoms together with a broad range of candidate female-specific symptoms derived from developmental research and clinical practice.

Longitudinal research in epidemiologic samples is required to evaluate whether female-specific diagnostic protocols would improve prediction of long-term prognosis. Follow-ups should assess a wide range of outcomes relevant to females (e.g., violence in intimate relationships, depression, reproductive and physical health), because women's prognostic outcomes extend beyond crime and antisocial personality disorder.

Evaluations in clinical settings are also required, to assess whether female-specific diagnostic protocols can improve prognosis prediction within clinical populations. Although evidence from longitudinal studies of community cohorts has not supported relaxing symptom thresholds for females, it is not known whether this finding translates into clinical settings where girls and boys are often referred for treatment for different reasons. Also, many treatment programs for adolescents with CD are sex-segregated and sex-tailored. Research should test whether female-specific diagnostic protocols would enhance the relevance of diagnosis for treatment planning (Levene et al., 2001).

To summarize, while there is currently no compelling evidence to include female-specific diagnostic protocols, additional research that systematically compares both sexes is required. While research on CD among girls has been growing rapidly, there is still a need to better understand the basic phenomenon of CD among girls (Hipwell et al., 2002; Pepler et al., 2005). However, female-only samples are not recommended because research must systematically compare both sexes within a cohort to evaluate whether female-specific diagnostic protocols are required.

A note on other group-specific diagnostic protocols

The question of female-specific diagnostic protocols is part of a much larger question: Should DSM and ICD allow diagnostic protocols tailored for patients of different sexes, ages, races, ethnic groups, cultures, religions, and even generations since immigration? There was vigorous debate about this at the February 2007 meeting of the Committee to Assess Research Needs for DSM-V Externalizing Disorders; here we summarize the major themes. Initial analyses using item-response theory (IRT) suggest that individual CD symptom criteria may carry different weights and meanings in different ethnic research samples in the United States. Thus, group-specific protocols might offer greater precision, and better clinical sensitivity to patient needs. However, support was also voiced for a universally applicable core diagnostic protocol. One objection to group-specific protocols was that the DSM and ICD should provide uniform criteria worldwide to prevent misuse of psychiatry for oppression of minority groups (e.g., using psychiatric drugs or hospitalization to silence political dissidents). It was also thought that, within

a nation, uniform criteria support the idea of fair and equal access to educational services and health-care regardless of sex, race, or culture. A pragmatic objection to group-specific protocols was that these could oblige diagnosticians to use complex algorithms. For example, many symptom weightings could apply if the patient were female, Black race, Hispanic ethnicity, and a third-generation immigrant from Puerto Rico to New York, or if the patient were male, Asian race, Muslim religious culture, and a recent new immigrant from Indonesia to Amsterdam. Moreover, in some regions undergoing rapid cultural change, research-based CD symptom weights derived on today's generation might not apply to tomorrow's young people. DSM-IV approaches this thorny issue of group-specific diagnosis by giving diagnosticians permission to discount a CD symptom if it represents a normative adaptation to the patient's social context (p. 88). Whether such trust in clinical judgment should be supplanted by some more formal arrangement must be subject to debate and research. Of key interest is whether subgroup-specific CD protocols derived from IRT analyses provide incremental predictive validity for patients in cultural subgroups, for example, by improving prediction of their prognosis.

Issue 9, early substance use as a criterion for CD?

Definition of the issue

A criterion for CD called 'substance abuse' (not to be confused with the formal DSM diagnosis of substance abuse) was included in DSM-III, but was subsequently dropped from DSM-III-R and DSM-IV. Drinking alcohol, smoking, and the use of illegal substances are currently listed in the DSM-IV as associated descriptive features for CD (p. 87). The question of this section is: should early substance use be considered a part of the diagnostic criteria for DSM-V CD or continue as an associated feature?

There appear to be two main reasons why early substance use was dropped from the DSM CD criteria. First, there was concern about the inclusion of status offenses that are pathological only when they occur among children. CD can be diagnosed among adults (at least in theory), and substance use would probably not be a valid indicator of CD among adults or older adolescents (Robins, 1987, 1991). Early substance use as a DSM-III criterion was too imprecise and there was no guidance about the age at which drinking modest amounts of alcohol no longer constitutes problematic substance use (Robins, 1987). Robins (1987) recommended the use of age limits for this criterion to resolve this problem, but the proposed age limits were never incorporated into the DSM. Second, although an item assessing early use of tobacco or drugs had adequate psychometric properties in the DSM-III-R field trials, it was

eliminated from the proposed diagnostic criteria for CD because it did not describe behavior that was inherently antisocial (Spitzer, Davies, & Barkley, 1990).

Given that there is disagreement among experts about excluding early substance use as a criterion for CD, the question is worth revisiting for DSM-V. Formally, early substance use is the precocious, age-inappropriate use of substances by children and adolescents (even when it would not meet the criteria for a substance-use disorder diagnosis). The challenge of including early substance use as a diagnostic criterion for CD in the DSM will be the development of a precise definition that incorporates appropriate age limits and identifies children who are involved in problematic behavior, and not merely casual experimentation or the developmentally-normative substance use of later adolescence. An example criterion might be something like 'Repeated use of alcohol, tobacco, or an illegal substance (without parental permission) prior to the age of 13.'

Rationale for adding early substance use to the criteria for CD for DSM-V

Substance use and CD are intimately connected (Robins, 1998), even at the level of individual events – that is, many adolescents commit antisocial acts while under the influence of alcohol and other drugs (White et al., 2002). Research has demonstrated that early substance use and antisocial behavior are indicators of an underlying 'problem behavior' dimension among adolescents (e.g., Donovan & Jessor, 1985), and that early alcohol use and CD share overlapping genetic and family-environment etiological factors (e.g., McGue et al., 2001). Thus, there may be a common underlying vulnerability predisposing children to use substances and to engage in other behaviors characteristic of CD.

Early substance use by children and adolescents is ignored in the DSM system, despite the fact that it can have dire short- and long-term consequences for youth (e.g., NIAAA, 2004/2005). Early substance use is a marker of risk for the later development of substance use disorders (e.g., Grant & Dawson, 1997, 1998), as well as associated disorders such as major depression and antisocial personality disorder (e.g., McGue & Iacono, 2005). There are three main reasons why the most logical home for this problematic behavior would be among the symptoms of CD. First, community-based epidemiologic studies consistently find that substance use among adolescents is more strongly associated with CD than with any other psychiatric disorder (Armstrong & Costello, 2002). Second, popular dimensional inventories of externalizing behavior problems typically include items related to substance use (e.g., Achenbach & Rescorla, 2001; Elliott, Huizinga, & Ageton, 1985). Third, purchasing alcohol, cigarettes, and illicit drugs, consuming alcohol underage, and having

illicit drugs in one's possession are all illegal for juveniles in most states in the United States. Thus, early substance use would fit into the CD symptom grouping of 'serious violation of rules' and would be consistent with this essential feature of CD.

In addition, although substance use and CD are relatively rare in pre-adolescent girls, there is evidence to suggest that disruptive behavior disorders in girls rarely occur in the absence of the use of alcohol, tobacco, or illicit drugs (Federman et al., 1997). Early substance use may be an especially sensitive indicator of CD among girls.

Predictive validity of early substance use in CD

Perhaps because substance use is one of the later developing symptoms of CD (e.g., Robins, 1987), there is limited research examining the extent to which early substance use prospectively predicts the escalation or persistence of CD. The few studies that have been done suggest that early substance use presages worse outcomes over time. For example, in a community-based study of male and female urban schoolchildren, children who had used alcohol without parental permission by ages 10–12 had higher initial levels of CD symptoms and more rapid increases in CD symptoms over the subsequent two years compared to 10–12-year-old children who had abstained from alcohol use (Johnson et al., 1995). In a study of clinic-referred boys, repeated marijuana use from ages 13–17 was the strongest predictor in a multivariate model of the progression from CD in adolescence to antisocial personality disorder in early adulthood (Loeber et al., 2002).

Reliability of assessing early substance use among children and adolescents

Substance use can be reliably assessed among adolescents (Winters et al., 2002). Substance use among elementary school-aged children is relatively under-studied, but there is evidence to suggest that this can be reliably assessed as well (Donovan et al., 2004). Reliably establishing the age of first substance use retrospectively among adolescents and adults might prove difficult due to telescoping, which is the tendency to recall onset of substance use as years later than it truly occurred (Parra et al., 2003). Like other covert antisocial behaviors that are symptoms of CD, it is essential to obtain self-reports directly from the young patient because parents often are unaware of their children's substance involvement.

Possible disadvantages of including early substance use as a CD criterion

Even if a criterion of early substance use were precisely defined to exclude experimentation and age-normative use, there might still be disadvantages. Including early substance use as a CD criterion will probably exacerbate the problem of comorbidity in the DSM because more individuals may meet diagnostic criteria for both CD and a substance use disorder. Including early substance use within the CD diagnosis may also have disadvantages for research. The availability of a relatively 'clean' CD construct facilitated much of the important research examining the relation between substance use-related behaviors and CD and in identifying shared and unique risk factors for these inter-related problem behaviors.

What research into early substance use and CD is needed?

The most critical research in evaluating whether early substance use should be included as a CD criterion in DSM-V will be in the area of criterion development. An age cut-off for 'early' must be determined and the wording of the symptom must effectively differentiate developmentally-appropriate substance use behaviors from behaviors indicative of CD. These goals can be accomplished by conducting in-depth surveys of substance involvement and psychopathology in representative community-based samples of children and adolescents, with retest assessments included to evaluate the reliability of candidate criteria. The prevalence of these new candidate criteria should be estimated as well as their effect on the prevalence of CD. Are more children and adolescents being diagnosed with CD with the inclusion of a new candidate criterion than without it? Comparisons of the prevalence of CD in boys and girls before and after including a new candidate criterion will also be informative. Does the inclusion of early substance use reduce the gender disparity in the diagnosis of CD? Finally, does a CD diagnosis including early substance use identify children who are impaired and in need of treatment better than the current CD diagnosis?

In the end, the decision of whether to include early substance use as a criterion for CD may rest upon the answers to questions that may not be completely answerable through empirical research. For example, when is a correlate of a disorder considered a core symptom versus an associated feature?

Issue 10, category or continuum?

Definition of the issue

DSM-IV takes a categorical approach to disorders, including CD. DSM-IV lists 15 symptom criteria for CD; children meeting three criteria receive the diagnosis. Recently, support has emerged for including dimensional operationalizations of all disorders in DSM-V, as a complementary option to accompany traditional diagnostic categories (First, 2006a; Helzer, Kraemer, & Krueger, 2006; Krueger, Watson,

& Barlow, 2005; Widiger & Samuel, 2005). Ahead of this wave, the continuum approach to assessing CD has enjoyed strong empirical support for many years (Achenbach, 1985; Robins, 1978).

Rationale for considering a continuum of CD for DSM-V

The categorical diagnosis of CD has a number of disadvantages that do not afflict dimensional approaches. First, with a categorical approach, variation in severity of dysfunction among children falling below and above the cut-off is lost (Hinshaw, Lahey, & Hart, 1993). Second, categories can create a false impression of change in a disorder's course when patients who slip below the cut-off by a symptom or two are considered to have recovered. In a longitudinal study, 80% of CD-diagnosed children evidenced apparent remission between repeated assessment intervals, but when CD symptoms were analyzed on a continuum, the cohort's mean-level and rank-order stability were moderate to strong across ten years, indicating little substantive change (Moffitt, Caspi, Rutter, & Silva, 2001). Third, unless there is evidence that conduct problems operate in a categorical fashion, the cut-off point inevitably must be a matter of convention. Studies searching for evidence of a categorical threshold point along the distribution of CD symptoms have not found it (e.g., Lahey et al., 1994). An epidemiological study compared potential cut-off points of 2, 3, 4, 5, and 6 symptom criteria, and reported that each was an arbitrary point along a linear continuum of CD severity (Moffitt et al., 2001). Arbitrary definitions could undermine confidence in the legitimacy of the DSM-V.

Predictive validity of the CD continuum

One prospective cohort study compared the predictive validity of the categorical diagnosis versus dimensional measurement of CD (Fergusson & Horwood, 1995). The dimensional variable was the better predictor of outcome. With increasing severity of CD along the continuum, there was increasing risk for juvenile offending and school dropout. Another cohort study reported a similar dose–response relationship in which the number of CD symptoms on a continuum (below and above the diagnostic cut-off) predicted later adult outcomes of education, work-life, relationships, parenting, mental health, physical health, substance abuse, and crime (Moffitt et al., 2001). Such dose–response patterns in follow-up studies indicate that variation in symptom severity both above and below the diagnostic cut-off is informative about prognosis.

Reliability of assessing CD as a continuum

In current research and clinical practice, CD is already commonly defined as both category and continuum. The most popular assessment tools reflect these two conceptualizations. Structured diagnostic interviews operationalize the specific DSM-IV CD criteria to achieve a categorical diagnosis (Costello et al., 1996; Goodman, Ford, Richards, Gatward, & Meltzer, 2000; Shaffer et al., 2000). Dimensional instruments cover a broad variety of conduct problems on symptom checklists (Achenbach & Rescorla 2001; Elander & Rutter 1996; Goodman, 1997). Dimensional instruments operate on two evidence-based principles: (a) aggregation of symptoms enhances reliability, and (b) the variety of antisocial behaviors is the best predictor of poor prognosis. Assessing CD as a continuum is widely acknowledged to be practically feasible and psychometrically sound (Koot, Crijnen, & Ferdinand, 1999).

Regarding the DSM-IV subtypes of childhood-onset versus adolescent-onset CD, two dimensional scales (called 'aggression' versus 'rule-breaking') have been shown to map fairly well respectively onto the two CD subtype constructs in research samples (Achenbach & Rescorla, 2001; Moffitt, 2006; Tackett, Krueger, Iacono, & McGue, 2005).

Possible disadvantages of a continuum approach to CD

In relation to other disorders, there is concern about the risk of including in DSM-V dimensional definitions that are 'useful for researchers but unfamiliar, burdensome, and of unknown utility to clinicians' (First, 2006a, p. 1679). However, assessing CD on a continuum is familiar to most pediatric clinicians, and it has known utility for predicting prognosis. DSM-IV implicitly acknowledges this by specifying CD severity (albeit imprecisely) as mild, moderate, or severe on the basis of symptom criteria in excess of the required three (p. 91). Clinicians will need diagnoses to guide the categorical decision to treat or not to treat. Nevertheless, complementing the DSM-V CD diagnosis by adding a more formal, structured continuum approach would not appear to have marked disadvantages.

What research on dimensional measures of CD is needed?

DSM-III improved psychiatric research and practice by encoding reliable, standardized definitions of diagnostic categories. Preparation of DSM-V is guided by a recognition that the scientific basis of psychiatry can now be further improved by adding reliable, standardized dimensional definitions of each disorder. For most mental disorders other than CD, there is a need for basic psychometric research to develop dimensional assessment tools, and epidemiological research to ascertain how the resulting dimensional scores relate to factors such as age, sex, ethnicity, developmental subtypes, and prognosis (First, 2006b). For these other disorders, psychometric

evaluations must be made to address questions about the best techniques for achieving dimensional measures. As examples: Should some symptoms receive heavier weighting than others, to reflect their greater severity? Must such weights differ at different ages? What response options should be used to record each symptom; should codes reflect simple presence versus absence (yes, no), or symptom frequency over time (never, sometimes, often), or severity (mild, moderate, severe)? Are patients' self-reports on symptom checklists a valid basis for making a diagnosis? What is the best way to combine ratings from different reporters, such as mothers and teachers? For CD, this psychometric work has by and large already been accomplished, much of it in preparation for DSM-IV (Frick et al., 1994). Conclusions about the best ways to measure the dimension of antisocial behavior have already been published in developmental psychopathology journals and criminology journals (most publications conclude that for CD, the simplest approaches are best). However, these findings have not been organized, and thus a meta-analysis to summarize the findings would be very useful before DSM-V.

Issue 11, life-course continuity from ODD to CD to ASPD?

Definition of the issue

This section addresses the degree of continuity and discontinuity in the association from ODD to CD to ASPD over the life-course. That is, are there empirical links among these disorders? And if so, do these links represent a pattern of heterotypic continuity that constitutes meaningful progression in the course of a single disorder, despite developmental changes in symptomatology? If these different disorders reflect heterotypic continuity in a single disorder, what is the core latent feature that unites them? Regrettably, thorough treatments of all nosological issues related to ODD or ASPD are beyond the scope of this article. We limit our focus to how ODD and ASPD relate developmentally to CD.

DSM-IV organizes ODD, CD, and ASPD hierarchically and developmentally, as if they reflect age-dependent expressions of the same underlying disorder. This hierarchical organization is reflected in the fact that DSM-IV does not allow for concurrent comorbidity among these disorders. ODD is conceptualized as a developmental precursor to CD. DSM-IV further states that 'all of the features of ODD are usually present in CD' (p. 93). ODD may be diagnosed only if criteria for CD are not also met. In turn, CD is similarly conceptualized as a developmental precursor to ASPD. A diagnosis of ASPD requires evidence of CD before age 15 years. For individuals over age 18 years, a diagnosis of CD can be given only if criteria for ASPD are not also met. Interestingly, whereas ODD and CD are Axis I

disorders in DSM-IV, ASPD is 'downgraded' hierarchically, to an Axis II disorder. Thus, if there is an underlying disorder uniting ODD, CD, and ASPD, a patient who has this one disorder would be inexplicably shifted from Axis I to Axis II on his or her 18th birthday.

Rationale for considering the continuity from ODD to CD to ASPD

The future of ODD and ASPD were discussed at the February 2007 meeting of the Committee to Assess Research Needs for DSM-V Externalizing Disorders; here we summarize the main themes. Regarding ASPD, preliminary consensus was reached for a recommendation to move it to Axis I, and re-label it Antisocial Disorder. Regarding ODD, some sentiment was voiced for reconsidering its status as a disorder. Critics of the DSM often point to ODD as an example of how psychiatry errs by defining normal behavior as pathological; critics argue that oppositional behavior is a reasonable part of the terrible-twos or teen-aged rebellion, not a mental illness. Some committee members thought ODD was transient and benign unless accompanied by ADHD, CD, or depression, implying that ODD might be redefined as a complicating condition for other disorders in DSM-V. Some members noted that ODD symptoms are virtually synonymous with the content of personality traits called high negative emotionality and low agreeableness (e.g., Lahey & Waldman, 2003), implying that ODD might be redefined as a personality dimension in DSM-V. Some members thought ODD involving conflict between a child and one parent could be redefined as a relational disorder in DSM-V, implying that ODD should be reserved for children who meet oppositional-defiant criteria in more than one setting, with a teacher as well as a parent. Proponents of ODD as a disorder noted that it is a precursor not only to CD, but also to ADHD, depression, anxiety, bipolar disorder, and substance abuse (e.g., Speltz et al., 1999); implying that treating ODD provides a valuable opportunity to prevent many other disorders. Proponents also noted that a diagnosis of ODD serves many families as a 'soft option,' promoting timely use of treatment services for a child, while avoiding a potentially more damaging diagnostic label. Each of the aforementioned points of view warrants research to build an evidence base. Here we ask how ODD relates to CD and ASPD.

Cross-sectional overlap between ODD, CD, and ASPD

This review did not identify any reports of cross-sectional overlap between CD and ASPD during adulthood, perhaps because most studies of adults have implemented the exclusionary rule in which CD is only diagnosed if ASPD is absent.

However, there is evidence of considerable construct overlap between ODD and CD from cross-sectional studies. In general, this evidence shows that some children with ODD also have CD features at the same age, whereas most children with CD also have ODD features at the same age. In two studies, research diagnostic criteria for ODD were relaxed so that CD did not exclude concurrent ODD. First, in an epidemiological sample of British youth, the percentage of children with ODD who also had CD increased with age, from about 10% among 5–7-year-olds to 60% among 13–15-year-olds (Maughan et al., 2004). The percentage of children with CD who also had ODD was about 60%, with no major variation by age. Second, in the Great Smoky Mountains Study community sample, about 30% of youth with CD met full criteria for concurrent ODD, but 95% of CD youth had at least one ODD symptom (Rowe, Maughan, Pickles, Costello, & Angold, 2002).

Predictive validity of ODD to CD and CD to ASPD

The hypothesis of continuity among ODD, CD, and ASPD has not been widely studied. Nevertheless, the available studies have used two longitudinal designs, follow-forward and follow-back. As a rule of thumb, the two designs reveal complementary pictures regarding continuity. Follow-back studies show that most CD children had prior ODD, and most (if not all) ASPD adults had prior CD. In contrast, follow-forward studies show that most ODD children do not develop CD, and most CD children do not develop ASPD. Thus, adult ASPD indicates a longstanding history of antisocial disorder from early life. However, children who begin life with an antisocial disorder need not progress toward ASPD. Indeed most such children recover (as noted in this article's section on CD subtypes).

Typically, ODD symptoms appear first, and then in a subset of ODD boys, CD symptoms follow (Lahey et al., 1997; Loeber et al., 1995). In a follow-back analysis in the Developmental Trends Study of clinic-referred boys, 80% of boys with childhood CD had prior ODD (Lahey et al., 1997). In a follow-forward analysis from this same sample, about 60% of all boys with ODD subsequently progressed to later CD (Lahey et al., 1997). A follow-forward analysis from the Great Smoky Mountains Study of a community sample showed that 40% of boys with ODD progressed to CD (Rowe et al., 2002).

Although clearly not all youth with CD progress toward an adult ASPD diagnosis, the degree of continuity is strong; follow-forward analyses show that about one-third to one-half of children with CD grow up to have adult ASPD (Robins, 1966, 1978). In the Developmental Trends Study's follow-forward analyses, 50% of youths with CD in adolescence later had ASPD (Loeber, Burke, & Lahey, 2002). Follow-back analyses show that even when ASPD is diagnosed for research purposes without requiring pre-existing CD symptoms, most ASPD individuals are found to have prior CD. In the Developmental Trends Study of a clinical cohort, 90% of adults with ASPD had a prior CD diagnosis (Loeber et al., 2002). In the Dunedin Longitudinal Study of a community cohort, 60% of adults with ASPD had a prior CD diagnosis (Kim-Cohen et al., 2003).

Many longitudinal studies report strong continuity from childhood to adulthood using dimensional measures of antisocial behavior or through trajectory analyses. However, almost no studies have examined the long-term continuity of diagnosed disorders in a longitudinal study from childhood ODD to adult ASPD. Existing evidence suggests that very few children who meet criteria for ODD in adolescence progress to ASPD without also having intermediate CD (Lahey, Loeber, Burke, & Applegate, 2005; Loeber et al., 2002). Also, some evidence suggests that individuals who do progress from ODD to CD to ASPD represent a more serious level of antisocial psychopathology, relative to children who recover and do not progress. For example, ODD is more often present in the history of the early-onset than the late-onset subtype of CD (Lahey et al., 1997), and ODD progresses to CD more often in boys than girls (Rowe et al., 2002). Also youths with CD who progress to ASPD, as compared to youths who do not progress to ASPD, more often exhibit callous-unemotional traits, comorbid depression, marijuana use, and serious violent behavior (Loeber et al., 2002). This evidence seems consistent with the findings about CD subtypes, which indicate that childhood-onset, life-course persistent antisocial behavior represents a more serious (and male-typical) form of antisocial psychopathology, as compared to the shorter-term childhood-limited and adolescent-onset subtypes of CD (see this article's section on subtypes).

What research on life-course continuity from ODD to CD to ASPD is needed?

First, research is needed on whether an individual's continuity among the three diagnoses constitutes a specific subtype of disorder unified by a core latent feature or features. Unpacking the broader diagnostic categories into a more specific subtype characterized by continuity across the life course may help identify whether a unifying psychopathological diathesis links ODD, CD, and ASPD (Krueger, 1999). Elsewhere in this article we have discussed leading candidates for this unifying core, including family history, biomarkers, and callous-unemotional traits. We have not focused on hyperactivity-impulsivity, but it too should be investigated as a leading candidate for the common core uniting ODD, CD, and ASPD. ADHD is highly comorbid with ODD (Egger & Angold, 2006). Several studies have pointed to ADHD as a developmental precursor of persistent, serious CD (Loeber et al., 1995; Mannuzza et al.,

2004; Moffitt et al., 1996). In particular, symptoms of hyperactivity and impulsivity predict early-onset CD (Loeber et al., 1995). Longitudinal follow-up studies are needed to distinguish whether ADHD might be a key syndrome that accounts for the subgroup of individuals who show continuity from ODD to CD to ASPD.

Second, the empirical support for shifting from Axis I (ODD, CD) to Axis II (ASPD) at age 18 should be evaluated (Hudziak et al., 2007). One potential key is whether the same personality trait abnormalities can be used to characterize individuals with ODD, CD, and ASPD. There is currently much enthusiasm for incorporating dimensional personality trait approaches to diagnostic constructs in DSM-V as a means of resolving oddities in the axis system (Krueger et al., 2005), and ASPD is frequently presented as a key example of the advantages of such personality trait approaches (First, 2006a).

Third, much previous research on ODD–CD–ASPD continuity has been conducted in males. What little is known about females suggests they show little continuity across time for the diagnosis of CD, because girls generally meet fewer criteria than boys (Moffitt, Caspi, Rutter, & Silva, 2001; Rowe et al., 2002). Research is needed to provide basic descriptive information about relations among ODD, CD, and ASPD in females.

Summary and conclusions

Since the DSM-IV appeared in 1994, an impressive amount of new information about CD has emerged. New biological correlates of CD have been discovered, resulting in a whole new world of intriguing physiological, neuroimaging, and genotype biomarkers associated with CD. Longitudinal birth cohorts launched in the 1970s and 1980s have now reached adulthood, and powerful new statistical procedures have been created to dissect the repeated measures from these cohorts, giving us our first views of conduct-problem trajectories from early childhood to mid-life. Groups formerly overlooked in CD research, such as girls and preschool children, have received vigorous research attention in the past 5 years, with provocative results. A concept formerly considered to be relevant to adults only, psychopathic callous-unemotional traits, has been successfully transported into the world of child and adolescent CD research. Progress in genetics research has recently revived enthusiasm about the potential of family-psychiatric-history data for understanding CD.

On one hand, these scientific advances change the way researchers and clinicians think about CD, and as a result, these advances generate some serious contenders for changing the diagnostic protocol for CD in DSM-V. On the other hand, our background work for this article indicated that the current DSM-IV CD protocol is widely considered to be very good, as it

is. We found no serious proposals to delete any of the current criteria for CD, or to shift the threshold for diagnosis up or down to 'correct' the prevalence rate of CD. Subtyping by age of onset enjoys a supportive evidence base, and although the age-of-onset system may need some tweaking, clinicians and researchers appear to be using it. In both research and clinical settings, using dimensional conduct-problem scales alongside the categorical CD diagnosis is already considered good practice, and reliable and valid dimensional scales are available. The current CD protocol involves a straightforward count of observable symptom behaviors; it is not much challenged by the complexities that engender controversy for other disorders (e.g., about core symptoms, requisite numbers of symptoms within different criterion sets, dubious subtypes, or blurred boundaries between strongly overlapping disorders). Overall, the current protocol for CD is short and simple, and it performs fairly well, at least in research settings. Finally, many pundits believe that DSM-V and ICD should maintain the status quo unless there is overwhelming evidence that obliges a change; even modest changes could transfigure patients' access to health-care and educational services, confuse the use of diagnoses in the courts, and undermine the cumulative nature of scientific research into mental disorders.

The contenders for change that we identified and reviewed here are mostly proposals to *add* something to the existing CD diagnostic protocol for DSM-V: a childhood-limited subtype, family psychiatric history, callous-unemotional traits, various and sundry biomarkers, female-specific criteria, preschool-specific criteria, or early substance use. Reasonable rationales have been put forward for adding each of these. However, in the absence of serious dissatisfaction with the current CD protocol, these contenders will have to present a far more compelling evidence base than is now available if they are to be considered for incorporation into DSM-V. Thus, we hope that vigorous efforts will be undertaken to meet the many research needs raised in this article. In particular, we found little evidence that biomarkers are ready to be incorporated into DSM-V, or even mentioned as associated features, because most biomarkers lack evidence of specificity to CD, evidence that they predict prognosis, and evidence that they can be translated for routine clinical use. We hope future biomarker research will take seriously our challenge to address questions of specificity, prediction, and translation.

To our surprise, although there is a great deal of interesting research into each of the 11 issues we reviewed in this article, very little of it has been strategically aimed toward providing the sort of evidence base that will be required to justify any alteration to the DSM-V. We found that many *case–control comparisons* have already been carried out; these have documented that each of the 11 issues

reviewed here is relevant to CD, and brought each to attention in the field. This article recommends other key designs that are fundamental for answering questions about whether or not an issue should be incorporated into DSM-V. *Epidemiological cohort* studies are needed to assess a diagnostic criterion's prevalence across age, sex, and race/ethnicity. Such cohort studies can also report the prevalence of 'abnormal' scores in the healthy population. If a criterion were added to the CD diagnosis, what rate of false positives would be expected? Cohort studies can also reveal whether the criterion in question is distributed as a category or a continuum in the population. *Psychiatric controls* are needed to evaluate the specificity of a diagnostic criterion to CD versus other disorders. *Longitudinal follow-up studies* are needed to test if a criterion improves prediction of CD children's course and prognosis. Such studies should follow up community cohorts, clinical samples, and forensic samples to insure that findings apply broadly. *Subtype comparisons* are needed because a criterion that seems to be only modestly related to CD children overall may in fact be strongly related to one specific subtype. *Translational research* is needed to convert researcher's data-collection paradigms into assessment tools that are practical in clinical and forensic settings. *Psychometric evaluations* are needed to assess the test–retest and inter-rater reliability of new assessment tools for CD. *Clinical trials* are needed to identify whether potential CD diagnostic criteria can predict treatment compliance or treatment response. These research approaches are urgently needed to prepare for DSM-V.

Acknowledgements

Correspondence to

Email: t.moffitt@iop.kcl.ac.uk

References

Ablow, J.C., Measelle, J.R., Kraemer, H.C., Harrington, R., Luby, J., Smider, N., et al., (1999). The MacArthur Three-City Outcome Study: Evaluating multi-informant measures of young children's symptomatology. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38, 1580–1590.

Achenbach, T.M. (1985). *Assessment and taxonomy of child and adolescent psychopathology*. Newbury Park, CA: Sage.

Achenbach, T.M., & Rescorla, L.A. (2001). *Manual for ASEBA School-Age Forms and Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

American Psychiatric Association. (1980). *Diagnostic and statistic manual of mental disorders* (3rd edn, DSM-III). Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd edn – rev., DSM-IIIR). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th edn, DSM-IV). Washington, DC: APA.

Andreasen, N.C., Endicott, J., Spitzer, R.L., & Winokur, G. (1977). Family history method using diagnostic criteria – reliability and validity. *Archives of General Psychiatry*, 34, 1229–1235.

Archer, J., & Coyne, S.M. (2005). An integrated review of indirect, relational and social aggression. *Personality and Social Psychology Review*, 9, 212–230.

Armstrong, T.D., & Costello, E.J. (2002). Community studies on adolescent substance use, abuse, or dependence and psychiatric comorbidity. *Journal of Consulting and Clinical Psychology*, 70, 1224–1239.

Arseneault, L., Kim-Cohen, J., Taylor, A, Moffitt, T.E., & Caspi, A. (2005). Psychometric evaluation of 5- and 7-year-old children's self-reports of conduct problems. *Journal of Abnormal Child Psychology*, 33, 537–550.

Arseneault, L., Moffitt, T.E., Caspi, A., Taylor, A., Rijsdijk, F.V., Jaffee, S.R., Ablow, J.C., & Measelle, J.R. (2003). Strong genetic effects on cross-situational antisocial behavior among 5-year-old children according to mothers, teachers, examiner-observers, and twins' self-reports. *Journal of Child Psychology and Psychiatry*, 44, 832–848.

Arseneault, L., Tremblay, R.E., Boulerice, B., & Saucier, J.-F. (2002). Obstetrical complications and adolescent violent behaviors: Testing two developmental pathways. *Child Development*, 73, 496–508.

Augimeri, L.K., Koegl, C.J., Webster, C.D., & Levene, K.S. (2001). *Early Assessment Risk List for Boys (EARL-20B)*. Toronto, Ontario: Earlscourt Child and Family Centre.

Bauer, L.O., & Hesselbrock, V.M. (2003). Brain maturation and subtypes of conduct disorder: Interactive effects on p300 amplitude and topography in male adolescents. *Journal of American Academy of Child and Adolescent Psychiatry*, 42, 106–115.

Beitchman, J.H., Baldassarra, L., Mik, H., De Luca, V., King, N., Bender, D., Ehtesham, S., & Kennedy, J.L. (2006). Serotonin transporter polymorphisms and persistent, pervasive childhood aggression. *American Journal of Psychiatry*, 163, 1103–1105.

Bennett, K.J., & Offord, D.R. (2001). Screening for conduct problems: Does the predictive accuracy of conduct disorder symptoms improve with age? *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 1418–1425.

Blair, R.J.R., Peschardt, K.S., Budhani, S., Mitchell, D.G.V., & Pine, D.S. (2006). The development of psychopathy. *Journal of Child Psychology and Psychiatry*, 47, 262–275.

Borum, R., Bartel, P., & Forth, A. (2006). *Manual for the Structured Assessment of Violence Risk in Youth (SAVRY)*. Lutz, FL: Psychological Assessment Resources, Inc.

Brennan, P.A., Grekin, E.R., & Mednick, S.A. (2003). Prenatal and perinatal influences on conduct disorder and delinquency. In B. Lahey, T.E. Moffitt, & A. Caspi, (Eds.), *Causes of conduct disorder and delinquency* (pp. 319–344). New York: Guilford Press.

Burke, J.D., Loeber, R., & Lahey, B.B. (2007). Adolescent conduct disorder and interpersonal callousness as predictors of psychopathy in young adults. *Journal of Clinical Child and Adolescent Psychology*, 36, 334–346.

Campbell, S.B. (2002). *Behavior problems in preschool children: Clinical and developmental Issues* (2nd edn). New York: Guilford.

Campbell, S.B., Spieker, S., Burchinal, M., & Poe, M.D. (2006). Trajectories of aggression from toddlerhood to age 9 predict academic and social functioning through age 12. *Journal of Child Psychology and Psychiatry*, 47, 791–800.

Cannon, M., Jones, P.B, & Murray, R.M. (2002). Obstetric complications and schizophrenia: Historical and meta-analytic review. *American Journal of Psychiatry*, 159, 1080–1092.

Caspi, A., Langley, K., Milne, B., Moffitt, T.E., O'Donovan, M., Owen, M.J., Polo Tomas, M., Poulton, R., Rutter, M., Taylor, A., Williams, B., & Thapar, A. (in press). A replicated molecular basis for subtyping antisocial behavior in ADHD. *Archives of General Psychiatry*.

Caspi, A., McClay, J., Moffitt, T.E., Mill, J., Martin, J., Craig, I.W., et al., (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, 297, 851–854.

Charney, D., Barlow, D., Botteron, K., Cohen, J., Goldman, D., Gur, R., Lin, K., Lopez, J., Meador-Woodruff, J., Moldin, S., Nestler, E., Watson, S., & Zalcman, S. (2002). Neuroscience research agenda to guide development of a pathophysiologically based classification system. In D.J. Rupfer, M.B. First & D.A., Regier (Eds.), *A research agenda for DSM-V* (pp. 31–83). Washington, DC: American Psychiatric Association.

Costa, L., Bauer, L., Kuperman, S., Porjesz, B., O'Connor, S., Hesselbrock, V., Rohrbaugh, J., & Begleiter, H. (2000). Frontal P300 decrements, alcohol dependence, and antisocial personality disorder. *Biological Psychiatry*, 47, 1064–1067.

Costello, E.J., Angold, A., Burns, B., Stangl, D., Tweed, D., & Erkanli, A. (1996). The Great Smokey Mountain Study of Youth, I: Prevalence and correlates of DSM-III-R disorders. *Archives of General Psychiatry*, 53, 1137–1143.

Coté, S., Vaillancourt, T., LeBlanc, J.C., Nagin, D.S., & Tremblay, R.E. (2006). The development of physical aggression from toddlerhood to pre-adolescence: A nationwide longitudinal study of Canadian children. *Journal of Abnormal Child Psychology*, 34, 71–85.

Coté, S., Zoccolillo, M., Tremblay, R.E., Nagin, D., & Vitaro, F. (2001). Predicting girls' conduct disorder in adolescence from childhood trajectories of disruptive behaviors. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 678–684.

Craddock, N., Owen, M.J., & O'Donovan, M.C. (2006). The catechol-O-methyl transferase (COMT) gene as a candidate for psychiatric phenotypes: Evidence and lessons. *Molecular Psychiatry*, 11, 446–458.

Crick, N.R., & Zahn-Waxler, C. (2003). The development of psychopathology in females and males: Current progress and future challenges. *Development and Psychopathology*, 15, 719–742.

Dadds, M.R., Perry, Y., Hawes, D.J., Merz, S., Riddell, A.C., Haines, D.J., Solak, E., & Abeygunawardane, A.I. (2006). Attention to the eyes and fear-recognition deficits in child psychopathy. *British Journal of Psychiatry*, 189, 280–281.

Deckel, A.W., Hesselbrock, V.M., & Bauer, L.O. (1996). Antisocial personality disorder, childhood delinquency, and frontal brain functioning: EEG and neuropsychological findings. *Journal of Clinical Psychology*, 52, 639–650.

Dick, D.M., Li, T-K., Edenberg, H.J., Hesselbrock, V., Kramer, J., Kuperman, S., Porjesz, B., Bucholz, K., Goate, A., Nurnberger, J. Jr., & Foroud, T. (2004). A genome-wide screen for genes influencing conduct disorder. *Molecular Psychiatry*, 9, 81–86.

Donovan, J.E., & Jessor, R. (1985). Structure of problem behavior in adolescence and young adulthood. *Journal of Consulting and Clinical Psychology*, 53, 890–904.

Donovan, J.E., Leech, S.L., Zucker, R.A., Loveland-Cherry, C.J., Jester, J.M., Fitzgerald, H.E., Puttler, L.I., Wong, M.M., & Looman, W.S. (2004). Really underage drinkers: Alcohol use among elementary students. *Alcoholism: Clinical and Experimental Research*, 28, 341–349.

Egger, H.L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry*, 47, 313–337.

Egger, H.L., Erkanli, A., Keeler, G., Potts, E., Walter, B.K., & Angold, A. (2006). Test–retest reliability of the Preschool Age Psychiatric Assessment (PAPA). *Journal of the American Academy of Child and Adolescent Psychiatry*, 45, 538–549.

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.

Elliott, D.S., Huizinga, D., & Ageton, S. (1985). *Explaining delinquency and drug use*. Beverly Hills, CA: Sage.

Farrington, D. (1997). The relationship between low resting heart rate and violence. In A. Raine, P.A. Brennan, D. Farrington, & S.A. Mednick (Eds.), *Biosocial bases of violence* (pp. 89–105). New York: Plenum.

Farrington, D.P., Gallagher, B., Morley, L., St. Ledger, R.J., & West, D. (1988). Are there any successful men from criminogenic backgrounds? *Psychiatry*, 51, 116–130.

Farrington, D.P., Jolliffe, D., Loeber, R., Stouthamer-Loeber, M., & Kalb, L.M. (2001). The concentration of offenders in families, and family criminality in the prediction of boys' delinquency. *Journal of Adolescence*, 24, 579–596.

Farrington, D.P., & Loeber, R. (2000). Epidemiology of juvenile violence. *Child and Adolescent Psychiatric Clinics of North America*, *9*, 733–748.

Federman, E.B., Costello, E.J., Angold, A., Farmer, E.M.Z., & Erkanli, A. (1997). Development of substance use and psychiatric comorbidity in an epidemiologic study of white and American Indian young adolescents: The Great Smoky Mountains Study. *Drug and Alcohol Dependence*, *44*, 69–78.

Fergusson, D.M., & Horwood, L.J. (1995). Predictive validity of categorically and dimensionally scored measures of disruptive childhood behaviors. *Journal of the American Academy of Child and Adolescent Psychiatry*, *34*, 477–487.

Fergusson, D.M., & Horwood, L.J. (2002). Male and female offending trajectories. *Development and Psychopathology*, *14*, 159–177.

First, M.B. (2006a). Beyond clinical utility: Broadening the DSM-V research appendix to include alternative diagnostic constructs. *American Journal of Psychiatry*, *163*, 1679–1681.

First, M.B. (2006b). Dimensional approaches in diagnostic classification: A critical appraisal. Conference Summary, June.

Forth, A.E., Kosson, D.S., & Hare, R.D. (2003). *The Psychopathy Checklist: Youth Version.* Toronto: Multi-Health Systems.

Frick, P.J., & Dickens, C. (2006). Current perspectives on conduct disorder. *Current Psychiatry Reports*, *8*, 59–72.

Frick, P.J., & Hare, R.D. (2001). *Antisocial process screening device.* Toronto: Multi-Health Systems.

Frick, P.J., Kimonis, E.R., Dandreaux, D.M., & Farrell, J.M. (2003). The 4 year stability of psychopathic traits in non-referred youth. *Behavioral Sciences and the Law*, *21*, 713–736.

Frick, P.J., Lahey, B.B., Applegate, B., Kerdyck, L., Ollendick, T., Hynd, G.W., Garfinkel, B., Greenhill, L., Biederman, J., Barkley, R.A., McBurnett, K., Newcorn, J., & Waldman, I. (1994). DSM-IV field trials for the disruptive behavior disorders: Symptom utility estimates. *Journal of the American Academy of Child and Adolescent Psychiatry*, *33*, 529–539.

Frick, P.J., & Marsee, M.A. (Eds.). (2006). *Psychopathy and developmental pathways to antisocial behavior in youth.* New York: Guilford.

Frick, P.J., O'Brien, B.S., Wootton, J.M., & McBurnett, K. (1994). Psychopathy and conduct problems in children. *Journal of Abnormal Psychology*, *103*, 700–707.

Frick, P.J., Stickle, T.R., Dandreaux, D.M., Farrell, J.M., & Kimonis, E.R. (2005). Callous-unemotional traits in predicting the severity and stability of conduct problems and delinquency. *Journal of Abnormal Child Psychology*, *33*, 471–487.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire (SDQ). *Journal of Child Psychology and Psychiatry*, *38*, 581–586.

Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The Development & Well-Being Assessment (DAWBA). *Journal of Child Psychology and Psychiatry*, *41*, 645–655.

Gorman-Smith, D., & Loeber, R. (2005). Are developmental pathways in disruptive behaviors the same for girls and boys? *Journal of Child and Family Studies*, *14*, 15–27.

Gould, T.D., & Gottesman, I.I. (2006). Psychiatric endophenotypes and the development of valid animal models. *Genes, Brain and Behavior*, *5*, 113–119.

Grant, B.F., & Dawson, D.A. (1997). Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence: Results from the National Longitudinal Alcohol Epidemiologic Survey. *Journal of Substance Abuse*, *9*, 103–110.

Grant, B.F., & Dawson, D.A. (1998). Age at onset of drug use and its association with DSM-IV drug abuse and dependence: Results from the National Longitudinal Alcohol Epidemiologic Survey. *Journal of Substance Abuse*, *10*, 163–173.

Haag, A.L. (2007). Biomarkers trump behavior in mental illness diagnosis. *Nature Medicine*, *13*, 3.

Haberstick, B.C., Smolen, A., & Hewitt, J.K. (2006). Family-based association test of the 5HTTLPR and aggressive behavior in a general population sample of children. *Biological Psychiatry*, *59*, 836–843.

Hall, M.H., Schulze, K., Rijsdijk, F., Picchioni, M., Ettinger, U., Bramon, E., Freedman, R., Murray, R.M., & Sham, P. (2006). Heritability and reliability of P300, P50 and duration mismatch negativity. *Behavior Genetics*, *36*, 845–857.

Hariri, A.R., & Holmes, A. (2006). Genetics of emotional regulation: The role of the serotonin transporter in neural function. *Trends in Cognitive Science*, *10*, 182–191.

Henry, B., Moffitt, T.E., et al., (1994). On the 'remembrance of things past': A longitudinal evaluation of the retrospective method. *Psychological Assessment*, *6*, 92–101.

Hawes, D.J., & Dadds, M.R. (2005). The treatment of conduct problems in children with callous-unemotional traits. *Journal of Consulting and Clinical Psychology*, *73*, 737–741.

Helzer, J.E., Kraemer, H.C., & Krueger, R.F. (2006). The feasibility and need for dimensional psychiatric diagnoses. *Psychological Medicine*, *36*, 1671–1680.

Hill, A.L., Degnan, K.A., Calkins, S.D., & Keane, S.P. (2006). Profiles of externalizing behavior in boys and girls across preschool: The roles of emotion regulation and inattention. *Developmental Psychology*, *42*, 913–928.

Hinshaw, S., Lahey, B.B., & Hart, E. (1993). Issues of taxonomy and comorbidity in the development of conduct disorder. *Development and Psychopathology*, *5*, 31–50.

Hipwell, A., & Loeber, R. (2006). Do we know which interventions are effective for disruptive and delinquent girls? *Clinical Child and Family Psychology Review*, *9*, 221–255.

Hipwell, A.E., Loeber, R., Stouthamer-Loeber, M., Keenan, K., White, H.R., & Krone-Man, L. (2002). Characteristics of girls with early onset disruptive and antisocial behavior. *Criminal Behavior and Mental Health*, *12*, 99–118.

Hipwell, A.E., Pardini, D.A., Loeber, R., Sembower, M., Keenan, K., & Stouthamar-Loeber, M. (2007). Callous-unemotional behaviors in young girls. Shared and unique effects relative to conduct problem. *Journal of Clinical Child and Adolescent Psychology*, *36*, 293–304.

Holmes, J., Payton, A., Barrett, J., Harrington, R., McGuffin, P., Owen, M., et al., (2002). Association of DRD4 in children with ADHD and comorbid conduct problems. *American Journal of Medical Genetics Part B*, *114*, 150–153.

Howell, J.C., & Hawkins, J.D. (1998). Prevention of youth violence. *Crime and Justice: A Review of Research*, *24*, 263–316.

Hudziak, J.J., Achenbach, T.M., Althoff, R.R., & Pine, D.S. (2007). A dimensional approach to developmental psychopathology. *International Journal of Methods in Psychiatric Research*, *16*, 16–23.

Huizinga, D., Haberstick, B.C., Smolen, A., Menard, S., Young, S.E., Corley, R.P., Stallings, M.C., Grotpeter, J., & Hewitt, J.K. (2006). Childhood maltreatment, subsequent antisocial behavior, and the role of monoamine oxidase A genotype. *Biological Psychiatry*, *60*, 677–683.

Hunt, S.C., Gwinn, M., & Adams, T.D. (2003). Family history assessment – strategies for prevention of cardiovascular disease. *American Journal of Preventive Medicine*, *24*, 136–142.

Ioannidis, J., Ntzani, E., Trikalinos, T., & Contopoulos-Ioannidis, D. (2001). Replication validity of genetic association studies. *Nature Genetics*, *29*, 306–309.

Jaffee, S.R., Belsky, J., Harrington, H., Caspi, A., & Moffitt, T.E. (2006). When parents have a history of conduct disorder: How is the caregiving environment affected? *Journal of Abnormal Psychology*, *115*, 309–319.

Johnson, E.O., Arria, A.M., Borges, G., Ialongo, N., & Anthony, J.C. (1995). The growth of conduct problem behaviors from middle childhood to early adolescence: Sex differences and the suspected influence of early alcohol use. *Journal of Studies on Alcohol*, *56*, 661–671.

Johnstone, T., Somerville, L.H., Alexander, A.L., Oakes, T.R., Davidson, R.J., Kalin, N.H., & Whalen, P.J. (2005). Stability of amygdala BOLD response to fearful faces over multiple scan sessions. *Neuroimage*, *25*, 1112–1123.

Kazdin, A.E., Holland, L, & Crowley, M. (1997). Family experience of barriers to treatment and premature termination from child therapy. *Journal of Consulting and Clinical Psychology*, *65*, 453–63.

Keenan, K., & Wakschlag, L.S. (2000). More than the terrible twos: The nature and severity of behavior problems in clinic-referred preschool children. *Journal of Abnormal Child Psychology*, *28*, 33–46.

Keenan, K., & Wakschlag, L.S. (2002). Can a valid diagnosis of disruptive behavior disorder be made in preschool children? *American Journal of Psychiatry*, *159*, 351–358.

Keenan, K., & Wakschlag, L.S. (2004). Are oppositional defiant and conduct disorder symptoms normative behaviors in preschoolers? A comparison of referred and nonreferred children. *American Journal of Psychiatry*, *161*, 356–358.

Keenan, K., Wakschlag, L.S., Danis, B., Hill, C., Humphries, M., Duax, J., & Donald, R. (2007). Further evidence of the reliability and validity of DSM-IV ODD and CD in preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*, 457–468.

Kendler, K. (2005). 'A gene for...': The nature of gene action in psychiatric genetics. *American Journal of Psychiatry*, *162*, 1243–1252.

Kendler, K.S., Kuo, P., Webb, B.T., Kalsi, G., Neale, M.C., Sullivan, P.F., Walsh, D., Patterson, D.G., Riley, B., & Prescott, C.A. (2006). A joint genomewide linkage analysis of symptoms of alcohol dependence and conduct disorder. *Alcoholism: Clinical and Experimental Research*, *30*, 1972–1977.

Kendler, K.S., & Roy, M.A. (1995). Validity of a diagnosis of lifetime major depression obtained by personal interview versus family history. *American Journal of Psychiatry*, *152*, 1608–1614.

Kim, M.S., Kim, J.J., & Kwon, J.S. (2001). Frontal P300 decrement and executive dysfunction in adolescents with conduct problems. *Child Psychiatry and Human Development*, *32*, 93–106.

Kim-Cohen, J., Arseneault, L., Caspi, A., Polo Tomas, M., Taylor, A., & Moffitt, T.E. (2005). Validity of DSM-IV Conduct Disorder in 41/2–5-year-old children: A longitudinal epidemiological study. *American Journal of Psychiatry*, *162*, 1108–1117.

Kim-Cohen, J., Caspi, A., Moffitt, T.E., Harrington, H.L., Milne, B., & Poulton, R. (2003). Prior juvenile diagnoses in adults with mental disorder: Developmental follow-back of a prospective-longitudinal cohort. *Archives of General Psychiatry*, *60*, 709–717.

Kim-Cohen, J., Caspi, A., Rutter, M., Tomas, M.P., & Moffitt, T.E. (2006). The caregiving environments provided to children by depressed mothers with or without an antisocial history. *American Journal of Psychiatry*, *163*, 1009–1018.

Kim-Cohen, J., Caspi, A., Taylor, A., Williams, B., Newcombe, R., Craig, I., & Moffitt, T.E. (2006). MAOA, maltreatment, and gene–environment interaction predicting children's mental health: New evidence and a meta-analysis. *Molecular Psychiatry*, *11*, 903–913.

Kim-Cohen, J., & The E-Risk Team (in revision). Predictive validity of DSM-IV conduct disorder diagnosis in 41/2–5-year-old children to age 10. *American Journal of Psychiatry*.

Kochanska, G., Gross, J.N., Lin, M.H., & Nichols, K.E. (2002). Guilt in young children: Development, determinants, and relations with a broader system of standards. *Child Development*, *73*, 461–482.

Koot, H., Crijnen, A., & Ferdinand, R. (1999). *Child psychiatric epidemiology*. Rotterdam, The Netherlands: van Gorcum.

Krueger, R. (1999). The structure of common mental disorders. *Archives of General Psychiatry*, *56*, 921–926.

Krueger, R.F., Watson, D., & Barlow, D.H. (2005). Toward a dimensionally based taxonomy of psychopathology. *Journal of Abnormal Psychology*, *114*, 491–493.

Kruesi, M.J., Casanova, M.F., Mannheim, G., & Johnson-Bilder, A. (2004). Reduced temporal lobe volume in early onset conduct disorder. *Psychiatry Research*, *132*, 1–11.

Lahey, B.B., Applegate, B., Barkley, R.A., Garfinkel, B., McBurnett, K., Kerdyk, L., Greenhill, L., Hynd, G.W., Frick, P.J., & Newcorn, J. (1994). DSM-IV field trials for ODD and CD in children and adolescents. *American Journal of Psychiatry*, *151*, 1163–1171.

Lahey, B.B., Loeber, R., Quay, H.C., Applegate, B., Shaffer, D., Waldman, I., et al., (1998). Validity of DSM-IV subtypes of conduct disorder based on age of onset. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*, 435–442.

Lahey, B.B., Loeber, R., Burke, J.D., & Applegate, B. (2005). Predicting future antisocial personality disorder in males from a clinical assessment in childhood. *Journal of Consulting and Clinical Psychology*, *73*, 389–399.

Lahey, B.B., Loeber, R., Quay, H.C., Frick, P.J., & Grimm, J. (1997). Oppositional defiant disorder and conduct disorder. In T.A. Widiger, A.J. Frances, H.A. Pincus, R. Ross, M.B. First, & W. Davis (Eds.), *DSM-IV sourcebook* (vol. 3, pp. 189–209). Washington, DC: American Psychiatric Association.

Lahey, B.B., Van Hulle, C.A., Waldman, I.D., Rodgers, J.L., D'Onofrio, B.M., Pedlow, S., Rathouz, P., & Keenan, K. (2006). Testing descriptive hypotheses regarding sex differences in the development of conduct problems and delinquency. *Journal of Abnormal Child Psychology*, *34*, 737–55.

Lahey, B.B., & Waldman, I.D. (2003). A developmental propensity model of the origins of conduct problems during childhood and adolescence. In B.B. Lahey, T.E. Moffitt & A. Caspi (Eds.), *Causes of conduct disorder and serious delinquency* (pp. 76–117). New York: Guilford Press.

Levene, K.S., Augimeri, L.K., Pepler, D.J., Walsh, M.M., Webster, C.D., & Koegl, C.J. (2001). *Early Assessment Risk List for Girls: EARL-21G*. Toronto, ON: Earlscourt Child and Family Centre.

Levene, K.S., Walsh, M.M., Augimeri, L.K., & Pepler, D.J. (2004). Linking identification and treatment of early risk factors for female delinquency. In M.M. Moretti, C.L. Odgers, & M.A. Jackson (Eds.), *Girls and aggression: Contributing factors and intervention principles* (pp. 147–63). New York: Kluwer Academic/ Plenum Publishers.

Loeber, R., Burke, J.D., & Lahey, B.B. (2002). What are the adolescent antecedents to antisocial personality disorder? *Criminal Behavior and Mental Health*, *12*, 24–36.

Loeber, R., Green, S.M., Keenan, K., & Lahey, B.B. (1995). Which boys will fare worse? Early predictors of the onset of conduct disorder in a six-year longitudinal study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *34*, 499–509.

Lynam, D.R. (1997). Pursuing the psychopath: Capturing the fledgling psychopath in a nomological net. *Journal of Abnormal Psychology*, *106*, 425–438.

Lynam, D.R., Caspi, A., Moffitt, T.E., Loeber, R., & Stouthamer-Loeber, M. (2007). Longitudinal evidence that psychopathy scores in early adolescence predict adult psychopathy. *Journal of Abnormal Psychology*, *116*, 155–165.

Lynam, D.R., & Gudonis, L. (2005). The development of psychopathy. *Annual Reviews of Clinical Psychology*, *1*, 381–407.

Lyoo, I.K., Lee, H.K., Jung, J.H., Noam, G.G., & Renshaw, P.F. (2002). White matter hyperintensities on magnetic resonance imaging of the brain in children with psychiatric disorders. *Comparative Psychiatry*, *43*, 361–368.

McCart, M.R., Priester, P.E., Davies, W.H., & Azen, R. (2006). Differential effectiveness of behavioral parent-training and cognitive-behavioral therapy for antisocial youth: A meta-analysis. *Journal of Abnormal Child Psychology*, *34*, 527–543.

McClellan, J.M., & Speltz, M.L. (2003). Psychiatric diagnosis in preschool children (letter). *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*, 127–128.

McGue, M., & Iacono, W.G. (2005). The association of early adolescent problem behavior with adult psychopathology. *American Journal of Psychiatry*, *162*, 1118–1124.

McGue, M., Iacono, W.G., Legrand, L.N., & Elkins, I. (2001). Origins and consequences of age at first drink. II. Familial risk and heritability. *Alcoholism: Clinical and Experimental Research*, *25*, 1166–1173.

Mannuzza, S., Klein, R.G., Abikoff, H., & Moulton, J.L. (2004). Significance of childhood conduct problems to later development of conduct disorder among children with ADHD: A prospective follow-up study. *Journal of Abnormal Child Psychology*, *32*, 565–573.

Marsee, M.A., Silverthorn, P., & Frick, P.J. (2005). The association of psychopathic traits with aggression and delinquency in non-referred boys and girls. *Behavioral Sciences and the Law*, *23*, 803–817.

Maughan, B., Rowe, R., Messer, J., Goodman, R., & Meltzer, H. (2004). Conduct disorder and oppositional defiant disorder in a national sample: Developmental epidemiology. *Journal of Child Psychology and Psychiatry*, *45*, 609–621.

Maughan, B., & Rutter, M. (2001), Antisocial children grown up. In J. Hill & B. Maughan (Eds.), *Conduct disorders In childhood and adolescence* (pp. 507–552). Cambridge: Cambridge University Press.

Measelle, J.R., Ablow, J.C., Cowen, P.A., & Cowen, C.P. (1998). Assessing young children's views of their academic, social, and emotional lives: An evaluation of the self-perception scales of the Berkeley Puppet Interview. *Child Development*, *69*, 1556–1576.

Merikangas, K.R. (2002). Implications of genetic epidemiology for classification. In J.E. Helzer & J.J. Hudziak (Eds.), *Defining psychopathology in the 21st century* (pp. 195–209). Washington, DC: American Psychiatric Publishing, Inc.

Messer, J., Goodman, R., Rowe, R., Meltzer, H., & Maughan, B. (2006). Preadolescent conduct problems in girls and boys. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*, 184–191.

Meyer-Lindenberg, A., Buckholtz, J.W., Kolachana, B., Hariri, A.R., Blasi, G., Wabnitz, A., Honea, R., Verchinski, B., Callicott, J., Egan, M., Mattay, V., & Weinberger, D.R. (2006). Neural mechanisms of genetic risk for impulsivity and violence in humans. *Proceedings of the National Academy of Sciences*, *103*, 6269–6274.

Meyer-Lindenberg, A., & Weinberger, D.R. (2006). Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nature Reviews Neuroscience*, *7*, 818–827.

Moffitt, T.E. (1993), Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674–701.

Moffitt, T.E. (2003). Life-course persistent and adolescence-limited antisocial behavior: A 10-year research

review and a research agenda. In B. Lahey, T.E. Moffitt, & A. Caspi (Eds.), *The causes of conduct disorder and serious juvenile delinquency* (pp. 49–75). New York: Guilford.

Moffitt, T.E. (2005a). Genetic and environmental influences on antisocial behaviors: Evidence from behavioral-genetic research. *Advances in Genetics*, *55*, 41–104.

Moffitt, T.E. (2005b). The new look of behavioral genetics in developmental psychopathology: Gene–environment interplay in antisocial behaviors. *Psychological Bulletin*, *131*, 533–54.

Moffitt, T.E. (2006). Life-course persistent versus adolescence-limited antisocial behavior. In D. Cicchetti & D.J. Cohen (Eds.), *Developmental psychopathology, 2nd edn, Vol. 3. Risk, disorder, and adaptation* (pp. 570–598). New York: Wiley.

Moffitt, T.E., Caspi, A., Dickson, N., Silva, P., & Stanton, W. (1996). Childhood-onset versus adolescent-onset antisocial conduct problems in males: Natural history from ages 3–18 years. *Development and Psychopathology*, *8*, 399–424.

Moffitt, T.E., Caspi, A., Harrington, H.L., & Milne, B. (2002), Males on the life- course-persistent and adolescence-limited antisocial pathways: Follow-up at age 26 years. *Development and Psychopathology*, *14*, 179–207.

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*. New York: Cambridge University Press.

Moffitt, T.E., Caspi, A., & Rutter, M. (2005). Strategy for investigating interaction between measured genes and measured environments. *Archives of General Psychiatry*, *62*, 473–481.

Moore, T.M., Scarpa, A., & Raine, A. (2002). A meta-analysis of serotonin metabolite 5-HIAA and antisocial behavior. *Aggressive Behavior*, *28*, 299–316.

Moretti, M.M., Odgers, C.L., & Jackson, M.A. (2004). *Girls and aggression: Contributing factors and intervention principles*. New York: Kluwer Academic/Plenum Publishers.

Muthén, B., & Muthén, L. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Newbury Park, CA: Sage Publications.

Nagin, D.S. (1995). *Group-based modeling of development*. Cambridge, MA: Harvard University Press.

Nagin, D.S., Farrington, D.P., & Moffitt, T.E. (1995). Life-course trajectories of different types of offenders. *Criminology*, *29*, 163–190.

Nelson, R.J. (2006). *Biology of aggression*. New York: Oxford University Press.

NIAAA. (20042005). Alcohol and development in youth – a multidisciplinary perspective. *Alcohol Research and Health*, *28*, 105–176.

NICHD Early Child Care Research Network. (2004). Trajectories of physical aggression from toddlerhood to middle childhood. *Monographs of the Society for Research in Child Development*, *69*, No. 278.

Nigg, J.T., & Huang-Pollock, C.L. (2003). An early-onset model of the role of executive functions and intelligence in CD. In B.B. Lahey, T.E. Moffitt, &

A. Caspi (Eds.), *Causes of conduct disorder and juvenile delinquency* (pp. 227–253). New York: The Guildford Press.

NIH State of the Science. (2004). Preventing violence and related health-risking behaviors in adolescents. An NIH Sate of the Science Consensus report: http://consensus.nih.gov/PREVIOUSSTATEMENTS.htm#2004YouthViolence.

Odgers, C., Caspi, A., Poulton, R., Harrington, H.L, Thomson, M., Broadbent, J., Dickson, N., Sears, M., Hancox, B., & Moffitt, T.E. (2007). Conduct problem subtypes predict differential adult health burden. *Archives of General Psychiatry*, *64*, 476–484.

Odgers, C.L., Milne, B., Caspi, A., Crump, R., Poulton, R., & Moffitt, T.E. (2007). Predicting prognosis for the conduct-problem boy: Can family history help? *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*, 1240–1249.

Odgers, C.L., Moffitt, T.E., Caspi, A., Broadbent, J.M., Dickson, N.P., Hancox, R., Harrington, H., Poulton, R., Sears, M.R., & Thomson, W.M. (in press). Female and male antisocial trajectories: From childhood origins to adult outcomes. *Development and Psychopathology*.

Odgers, C.L., & Moretti, M.M. (2002). Aggressive and antisocial girls: Research update and challenges. *International Journal of Forensic Mental Health*, *1*, 103–19.

Odgers, C.L., Reppucci, N.D., & Moretti, M.M. (2005). Nipping psychopathy in the bud: An examination of the convergent, predictive, and theoretical utility of the PCL-YV among adolescent girls. *Behavioral Sciences and the Law*, *23*, 743–63.

Ortiz, J., & Raine, A. (2004). Heart rate and antisocial behavior in children and adolescents: A meta-analysis. *Journal of the American Academy of Child and Adolescents Psychiatry*, *43*, 154–162.

Orvaschel, H., & Puig-Antich, J. (1995). *Schedule for Affective Disorders and Schizophrenia for School-Age Children – Epidemiologic 5th Version*. Ft. Lauderdale, FL: Nova University.

Oxford, M., Cavell, T.A., & Hughes, J.N. (2003). Callous-unemotional traits moderate the relation between ineffective parenting and child externalizing problems: A partial replication and extension. *Journal of Clinical Child and Adolescent Psychology*, *32*, 577–585.

Pardini, D.A., Lochman, J.E., & Frick, P.J. (2003). Callous/unemotional traits and social-cognitive processes in adjudicated youths. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*, 364–371.

Parra, G.R., O'Neill, S.E., & Sher, K.J. (2003). Reliability of self-reported age of onset of substance involvement onset. *Psychology of Addictive Behaviors*, *17*, 211–218.

Penney, S.R., & Moretti, M.M. (2006). The relation of psychopathy to concurrent aggression and antisocial behavior in high-risk adolescent girls and boys. *Behavioral Sciences and the Law* [Epub ahead of print].

Pepler, D.J., Madsen, K.C., Webster, C., & Levene, K.S. (2005). *The development and treatment of girlhood aggression*. Mahwah, NJ: Lawrence Erlbaum Associates.

Phan, K.L., Wager, T., Taylor, S.F., & Liberzon, I. (2002). Functional neuroimaging studies of human emotions. *Neuroimage*, 16, 331–348.

Phillips, K.A., First, M.B., & Pincus, H.A. (2005). *Advancing DSM*. Washington, DC: American Psychiatric Association.

Piquero, A. (2005). *Taking stock of developmental trajectories of criminal activity over the life course.* Washington, DC: National Institute of Justice.

Piquero, A.R., Gibson, C.L., Daigle, L.E., Piquero, N., & Tibbetts, S.G. (2007). Are life-course-persistent offenders at risk for adverse health outcomes? *Research in Crime and Delinquency, 44,* 185–207.

Piquero, A., & Tibbetts, S. (1999). The impact of pre/perinatal disturbances and disadvantaged familial environment in predicting criminal offending. *Studies on Crime and Crime Prevention, 8,* 52–70.

Pompanon, F., Bonin, A., Bellemain, X., & Teberlet, P. (2005). Genotyping errors: Causes, consequences, and solutions. *Nature Reviews Genetics, 6,* 847–859.

Popma, A., & Raine, A. (2006). Will future forensic assessment be neurobiologic? *Child and Adolescent Psychiatric Clinics of North America, 15,* 429–444.

Putallaz, M., & Bierman, K.L. (2004). *Aggression, antisocial behavior, and violence among girls: A developmental perspective.* New York: Guilford Press.

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

Raine, A., Brennan, P., & Mednick, S.A. (1994). Birth complications combined with early maternal rejection at age 1 predispose to violent crime at age 18 years. *Archives of General Psychiatry, 51,* 984–988.

Raine, A., Venables, P.H., & Williams, M. (1990). Relationships between N1, P300, and contingent negative variation recorded at age 15 and criminal behavior at age 24. *Psychophysiology, 27,* 567–74.

Rhee, S.H., & Waldman, I.D. (2002). Genetic and environmental influences on antisocial behavior: A meta-analysis of twin and adoption studies. *Psychological Bulletin, 128,* 490–529.

Roberts, B.W., & DelVecchio, W.F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126,* 25–30.

Robins, L.N. (1966). *Deviant children grown-up: A sociological and psychiatric study of sociopathic personalities.* Baltimore: Williams and Wilkins.

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

Robins, L.N. (1987). An evaluation of the DSM-III diagnosis of conduct disorder. In G.L. Tischler (Ed.), *Diagnosis and classification in psychiatry: A critical appraisal of DSM-III* (pp. 384–395). Cambridge: Cambridge University Press.

Robins, L.N. (1991). Conduct disorder. *Journal of Child Psychology and Psychiatry, 32,* 193–212.

Robins, L.N. (1998). The intimate connection between antisocial personality and substance abuse. *Social Psychiatry and Psychiatric Epidemiology, 33,* 393–399.

Rowe, R., Maughan, B., Pickles, A., Costello, E.J., & Angold, A. (2002). The relationship between DSM-IV oppositional defiant disorder and conduct disorder: Findings from the Great Smoky Mountains Study. *Journal of Child Psychology and Psychiatry, 43,* 365–373.

Rutter, M., Kim-Cohen, J., & Maughan, B. (2006). Continuities and discontinuities in psychopathology between childhood and adult life. *Journal of Child Psychology and Psychiatry, 47,* 276–95.

Sakai, J.T., Young, S.E., Stallings, M.C., Timberlake, D., Smolen, A., Stetler, G.L., & Crowley, T.J. (2007). Case–control and within-family tests for an association between conduct disorder and 5HTTLPR. *American Journal of Medical Genetics, Part B, 141,* 825–832.

Schaeffer, C.M., Petras, H., Ialongo, N., Masyn, K.E., Hubbard, S., Poduska, J., & Kellam, S. (2006). A comparison of girls' and boys' aggressive-disruptive behavior trajectories across elementary school: Prediction to young adult antisocial outcomes. *Journal of Consulting and Clinical Psychology, 74,* 500–10.

Scott, E.S., & Grisso, T. (1997). The evolution of adolescence: A developmental perspective on juvenile justice reform. *The Journal of Criminal Law and Criminology, 88,* 137–189.

Sengupta, S.M., Grizenko, N., Schmitz, N., Schwartz, G., Ben Amor, L., Bellingham, J., de Guzman, R., Polotskaia, A., Ter Stepanian, M., Thakur, G., & Joober, R. (2006). COMT Val108/158Met gene variant, birth weight, and conduct disorder in children with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry, 45,* 1363–1369.

Shaffer, D., Fisher, P., Lucas, C.P., Dulcan, M.K., & Schwab-Stone, M.E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DSIC-IV). *Journal of the American Academy of Child and Adolescent Psychiatry, 39,* 28–38.

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

Simon, G.E., & VonKorff, M. (1995), Recall of psychiatric history in cross-sectional surveys: Implications for epidemiologic research. *Epidemiologic Reviews, 17,* 221–227.

Speltz, M.L., McClellan, J., DeKlyen, M., & Jones, K. (1999). Preschool boys with oppositional defiant disorder: Clinical presentations and diagnostic change. *Journal of the American Academy of Child and Adolescent Psychiatry, 38,* 838–845.

Spitzer, R.L., Davies, M., & Barkley, R.A. (1990). The DSM-III-R field trial of disruptive behavior disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 690–697.

Stadler, C., Sterzer, P., Schmeck, K., Krebs, A., Kleinschmidt, A., & Poustka, F. (2006). Reduced anterior cingulate activation in aggressive children and adolescents during affective stimulation: Association with temperament traits. *Journal of Psychiatric Research,* Mar 2; [Epub ahead of print].

Stallings, M.C., Corley, R.P., Dennehey, B., Hewitt, J.K., Krauter, K.S., Lessem, J.M., Milulich-Gilbertson, S.K., Rhee, S.H., Smolen, A., Young, S.E., & Crowley, T.J. (2005). A genome-wide search for quantitative trait loci that influence antisocial drug dependence in adolescence. *Archives of General Psychiatry, 62,* 1042–1051.

Sterzer, P., Stadler, C., Krebs, A., Kleinschmidt, A., & Poustka, F. (2005). Abnormal neural responses to emotional visual stimuli in adolescents with conduct disorder. *Biological Psychiatry*, *57*, 7–15.

Susman, E.J. (2006). Psychobiology of persistent antisocial behavior: Stress, early vulnerabilities and the attenuation hypothesis. *Neuroscience and Biobehavioral Reviews*, *30*, 376–389.

Tackett, J.L., Krueger, R.F., Iacono, W.G., & McGue, M. (2005). Symptom-based subfactors of DSM-defined conduct disorder: Evidence for etiologic distinctions. *Journal of Abnormal Psychology*, *114*, 483–487.

Task Force on Research Diagnostic Criteria: Infancy and Preschool. (2003). Research diagnostic criteria for infants and preschool children: The process and empirical support. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*, 1504–1512.

Taylor, J., Iacono, W.G., & McGue, M. (2000). Evidence for a genetic etiology for early-onset delinquency. *Journal of Abnormal Psychology*, *109*, 634–643.

Taylor, A., & Kim-Cohen, J. (2007). Meta-analysis of gene–environment interactions in developmental psychopathology. *Development and Psychopathology*, *19*, 1029–1037.

Thapar, A., Langley, K., Fowler, T., Rice, F., Turic, D., Whittinger, N., Aggleton, J., Ven den Bree, M., Owen, M., & O'Donovan, M. (2005). Catechol O-methyltransferase gene variant and birth weight predict early-onset antisocial behavior in children with attention-deficit/hyperactivity disorder. *Archives of General Psychiatry*, *62*, 1275–1278.

Thapar, A., O'Donovan, M., & Owen, M.J. (2005). The genetics of attention deficit hyperactivity disorder. *Human Molecular Genetics*, *14*, Spec No. 2, R275–282.

Thompson, W.D., Orvaschel, H., Prusoff, B.A., & Kidd, K.K. (1982). An evaluation of the family history method for ascertaining psychiatric-disorders. *Archives of General Psychiatry*, *39*, 53–8.

Tremblay, R.E. (2000), The development of aggressive behavior during childhood: What have we learned in the past century? *International Journal of Behavioral Development*, *24*, 129–141.

Tremblay, R.E. (2003). Why socialization fails. In B. Lahey, T.E. Moffitt, & A. Caspi (Eds.), *Causes of conduct disorder and delinquency* (pp. 182–226). New York: Guilford Press.

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

Van Goozen, S., Fairchild, G., Snoek, H., & Harold, G.T. (2007). The evidence for a neurobiological model of childhood antisocial behavior. *Psychological Bulletin*, *133*, 149–182.

Viding, E., Blair, R.J.R., Moffitt, T.E., & Plomin, R. (2005). Evidence for substantial genetic risk for psychopathy in 7-year-olds. *Journal of Child Psychology and Psychiatry*, *46*, 592–597.

Viding, E., & Blakemore, S-J. (2006). Endophenotype approach to the study of developmental disorders: Implications for autism research. *Behavior Genetics*, Sept [Epub ahead of print].

Wakschlag, L.S., Leventhal, B.L., & Thomas, J.M. (in press). Disruptive behavior disorders and ADHD in preschool children: Characterizing heterotypic continuities for a developmentally-informed nosology for DSM-V. In: D. Regier, M. First, & W. Narrow (eds.), *Age and gender considerations in psychiatric diagnosis: A research agenda for DSM-V*. Washington, DC: American Psychiatric Publishing, Inc.

Wiesner, M., Kim, H.K., & Capaldi, D. (2005). Developmental trajectories of offending: Validation and prediction to young adult alcohol use, drug use, and depressive symptoms. *Development and Psychopathology*, *17*, 251–270.

Weissman, M.M., Wickramaratne, P., Adams, P., Wolk, S., Verdeli, H., & Olfson, M. (2000). Brief screening for family psychiatric history: The family history screen. *Archives of General Psychiatry*, *57*, 675–82.

White, H.R., Tice, P.C., Loeber, R., & Stouthamer-Loeber, M. (2002). Illegal acts committed by adolescents under the influence of alcohol and drugs. *Journal of Research in Crime and Delinquency*, *39*, 131–152.

Widiger, T., & Samuel, D. (2005). Diagnostic categories or dimensions: A question for the Diagnostic and Statistical Manual of Mental Disorders-5th edn. *Journal of Abnormal Psychology*, *114*, 494–504.

Widom, C.S., & Brzustowicz, L.M. (2006). MAOA and the 'cycle of violence:' Childhood abuse and neglect, MAOA genotype, and risk for violent and antisocial behavior. *Biological Psychiatry*, *60*, 684–689.

Wiesner, M., & Capaldi, D.M. (2003). Relations of childhood and adolescent factors to offending trajectories of young men. *Journal of Research in Crime and Delinquency*, *40*, 231–262.

Winters, K.C., Latimer, W.W., & Stinchfield, R. (2002). Clinical issues in the assessment of adolescent alcohol and other drug use. *Behavior Research and Therapy*, *40*, 1443–1456.

Wootton, J.M., Frick, P.J., Shelton, K.K., & Silverthorn, P. (1997). Ineffective parenting and childhood conduct problems: The moderating role of callous-unemotional traits. *Journal of Consulting and Clinical Psychology*, *65*, 301–308.

Xie, H., Swift, D.J., Cairns, B., & Cairns, R.B. (2002). Aggressive behaviors in social interaction and developmental adaptation: A narrative analysis of interpersonal conflicts during early adolescence. *Social Development*, *11*, 205–224.

Yoon, P.W., Scheuner, M.T., & Khoury, M. (2003). Research priorities for evaluating family history in the prevention of common chronic diseases. *American Journal of Preventive Medicine*, *24*, 128–35.

Young, S., Smolen, A., Hewitt, J., Haberstick, B., Stallings, M., Corley, R., et al., (2006). Interaction between MAO-A genotype and maltreatment in risk for conduct disorder: Failure to confirm in adolescent patients. *American Journal of Psychiatry*, *163*, 1019–1025.

Zalecki, C.A., & Hinshaw, S.P. (2004). Overt and relational aggression in girls with attention deficit hyperactivity disorder. *Journal of Child and Adolescent Psychology*, *33*, 125–137.

Zimmerman, M., McGlinchey, J.B., Young, D., & Chelminski, I. (2006). Diagnosing major depressive disorder VII – family history as a diagnostic criterion. *Journal of Nervous and Mental Disease*, *194*, 704–7.

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

Zoccolillo, M., Tremblay, R., & Vitaro, F. (1996). DSM-III-R and DSM-III criteria for conduct disorder in preadolescent girls: Specific but insensitive. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 461–470.

Manuscript accepted 2 July 2007