Building an Evidence Base for *DSM–5* Conceptualizations of Oppositional Defiant Disorder and Conduct Disorder: Introduction to the Special Section

Dustin A. Pardini
University of Pittsburgh Medical Center

Paul J. Frick
University of New Orleans

Terrie E. Moffitt
Duke University and King’s College London

The *DSM–5* ADHD and Disruptive Behavior Disorders Work Group recently outlined a research agenda designed to support possible revisions to the diagnostic criteria for oppositional defiant disorder (ODD) and conduct disorder (CD). Some of the areas in need of further investigation include (a) examining the clinical utility of the current diagnostic system in girls, (b) further clarifying the developmental progression from ODD to CD, (c) determining whether facets of ODD symptoms can help explain heterotypic continuity and enhance predictive validity, (d) evaluating the clinical utility of a new subtyping scheme for CD on the basis of the presence of callous–unemotional traits, and (e) comparing the clinical utility of dimensional versus categorical conceptualizations of ODD and CD. This special section was organized in an attempt to provide data on these issues using a diverse array of longitudinal data sets consisting of both epidemiological and clinic-based samples that collectively cover a large developmental span ranging from childhood through early adulthood.

*Keywords:* callous-unemotional traits, oppositional defiant disorder, conduct disorder, longitudinal, *DSM–5*

The diagnosis of oppositional defiant disorder (ODD) refers to a persistent pattern of negativistic, hostile, defiant, and disobedient behaviors toward others, whereas conduct disorder (CD) is characterized by a persistent pattern of behavior that involves significant violations of the rights of others and/or major societal norms (American Psychiatric Association, 1968). Three theoretically distinct diagnoses were described as arising from divergent environmental factors: runaway reaction, unsocialized aggressive reaction, and group delinquent reaction. Rather than relying on specific diagnostic criteria, two- to three-sentence descriptions of each subtype were provided in the *DSM–II*, and it was up to the clinician to...
decide whether an individual’s behavior was consistent with one of the subtypes. The runaway subtype described individuals who fled threatening situations at home and engaged in covert acts of stealing as a means of self-preservation. The unsocialized aggressive subtype described loners with a pattern of hostile disobedience, aggressiveness, stealing, and lying whose behaviors arose primarily as a reaction to inconsistent discipline and parental rejection. The group delinquent reaction subtype included individuals who committed predominately nonaggressive acts as part of a delinquent peer group, and their behavior problems were purportedly a consequence of being poorly monitored in an impoverished neighborhood. These formulations for the subtypes originated from clinical observations as well as cluster and factor analysis studies with delinquent boys (Jenkins & Boyer, 1967; Quay, 1964).

The diagnosis of ODD (originally referred to as oppositional disorder) and CD was first introduced as part of the transition from DSM–II to the Diagnostic and Statistical Manual of Mental Disorders (3rd ed.; DSM–III; American Psychiatric Association, 1980). The DSM–III represented a fundamental paradigm shift in the formulation of all mental disorders, because specific diagnostic criteria were provided for each disorder in an attempt to address problems of unreliability in the DSM–II classification system (Rutter & Shaffer, 1980). The original diagnosis of ODD required that individuals exhibit at least two of five symptoms over the past 6 months, including violations of minor rules, temper tantrums, argumentativeness, provocative behavior, and stubbornness (American Psychiatric Association, 1980). In contrast, CD described a more severe form of deviant behavior that involved significant violations of the rights of others and/or major societal norms. Starting with the DSM–III, youths who met criteria for CD could not be diagnosed with ODD, because it was believed that ODD would already be present in nearly all individuals with CD (American Psychiatric Association, 1987). Four primary subtypes of CD were proposed on the basis of whether the individual exhibited aggressive (e.g., assault, rape) or nonaggressive (e.g., lying, stealing) conduct problems and whether the individual was socialized (e.g., has lasting friendships, feels guilt/remorse) or unsocialized (e.g., no close friendships, lacks guilt/remorse). These subtypes were based largely on the subtypes originally described in the DSM–II, and symptoms and diagnostic thresholds were specified for each subgroup.

Although empirical studies conducted prior to formulation of the DSM–III had supported the distinction between features of disruptive behavior disorders and affective disorders in children, there was little empirical evidence supporting the reliability and validity of making finer grained distinctions between ODD and CD and the subtypes of CD (Achenbach, 1980; Rutter & Shaffer, 1980). In fact, the DSM–III field trials indicated that the new symptom-based classification system produced only modest increases in interrater reliability for many childhood and adolescent disorders in comparison to the DSM–II (Mattison, Cantwell, Russell, & Will, 1979). The addition of CD as an independent psychiatric disorder was also severely criticized due to perceptions that it pathologized normative childhood behaviors (Rutter & Shaffer, 1980), especially in the absence of aggressive CD symptoms (Achenbach, 1980). The lack of empirical evidence supporting the proposed symptom thresholds for ODD and CD, as well as other mental disorders, also added fuel to the controversy regarding the DSM–III diagnostic system.

Due to the considerable criticism of DSM–III, a work group was formed 3 years after its publication to work on what was supposed to be a relatively minor revision (Rey, 1988). Although research on the legitimacy of the new ODD diagnosis was still sparse, evidence suggested that ODD was associated with only minor levels of impairment and was hard to reliably differentiate from CD (Rey et al., 1988). To address these issues, the disruptive and behavior disorders work group added several new ODD symptoms (e.g., spiteful or vindictive, angry and resentful) and clarified others (Spitzer, Davies, Barkley, & Costello, 1990). There were also explicit attempts to differentiate ODD symptoms from normative behavior. Specifically, symptoms had to occur more frequently than what is considered typical given the child’s age or developmental level to meet diagnostic threshold, although guidelines for what constituted “typical” were based on the clinician’s judgment (American Psychiatric Association, 1987). The diagnostic criteria for CD were also simplified to include a single set of symptoms (e.g., being physically cruel to people, destroying property, stealing, being truant) rather than four different subtypes. Each symptom needed to be present for at least 6 months to reach diagnostic threshold (Spitzer et al., 1990). Although this was done in part to address concerns about the reliability and validity of DSM–III CD subtypes, once individuals were diagnosed with CD they could still be placed into one of two primary subgroups on the basis of whether their deviant behaviors occurred primarily with peers (i.e., group type) or while alone (i.e., solitary aggressive type). This subtyping scheme had been supported by at least one longitudinal study that indicated that solitary aggressive delinquents were more likely to be convicted and incarcerated as adults compared with socialized delinquents (Henn, Bardwell, & Jenkins, 1980).

To empirically evaluate the proposed criteria for ODD and CD according to the Diagnostic and Statistical Manual of Mental Disorders (3rd ed., rev.; DSM–III–R; American Psychiatric Association, 1987), field trials were initiated within several mental health clinics (Spitzer et al., 1990). The field trial sample consisted of a total of 550 predominately male, school-aged children. The DSM committee decided that the sole empirical criterion that would be used to establish the validity of the newly proposed ODD/CD symptoms would be individual clinician diagnosis. Specifically, clinicians rated the proposed DSM–III–R symptoms in each child and then diagnosed the child on the basis of their expert clinical judgment using all available information, including symptoms and family factors not listed in the DSM (Spitzer et al., 1990). This reliance on individual clinical judgment as the gold standard in the field trials was controversial, given the known problems of unreliability and poor discriminant validity when relying on individual clinical opinion (Rutter & Shaffer, 1980; Werry, Methren, Fitzpatrick, & Dixon, 1983). Other acknowledged limitations of the field trials were the inclusion of a relatively small number of girls, the scarcity of preschool and adolescent-aged youths, and the reliance on a sample that included a minimal number of children diagnosed as having ODD or CD (Spitzer et al., 1990).

Analysis indicated that all of the newly proposed symptoms for ODD and CD had considerable discriminating power when using individual clinician diagnosis as the criterion. However, one proposed ODD symptom (i.e., bullying/being mean to others) and four proposed CD symptoms (i.e., having early sex, using drugs, bor-
rowing without permission, cheating in games) were eliminated by
the committee primarily on theoretical grounds (Spitzer et al., 1990). Symptom thresholds for each disorder were decided solely
on analysis of the sensitivity, specificity, and total predictive
power associated with different cutoff points, again with individual
clinician diagnosis as the criterion. In the end, it was determined
that an ODD diagnosis required at least five out of the nine new
symptoms to have been present over the past 6 months to reach
diagnostic threshold, whereas a diagnosis of CD required three out
of 13 newly defined symptoms to have been present for at least 6
months (American Psychiatric Association, 1987). However, the
inherent variability within diagnostic categories was recognized
with the initiation of severity classifications (i.e., mild, moderate,
severe) based on both the number of symptoms and the amount of
functional impairment present. The committee also decided to
retain the convention that CD preempted an ODD diagnosis be-
cause only 16% of children diagnosed with CD in the field trial
sample failed to meet criteria for ODD using the new symptom
thresholds (Spitzer et al., 1990). Although the committee acknowl-
edged that the observed overlap between ODD and CD was likely
less pronounced in community samples, they felt that building a
taxonomy of childhood disorders based on epidemiological sam-
ples did not provide any distinct advantages to the clinic-based
taxonomy of childhood disorders based on epidemiological sam-
pling method of the field trials (Spitzer et al., 1990).

With the publication of the DSM–III–R in 1987 (American
Psychiatric Association), there was a general acknowledgment that
the evidence base justifying the substantive revisions to ODD and
CD was slightly improved over previous iterations yet still lacking.
As such, the newly developed ODD/CD symptoms and thresholds
were viewed as a useful foundation for future refinements in
developing the Diagnostic and Statistical Manual of Mental Dis-
orders (4th ed.; DSM–IV; American Psychiatric Association, 2000; see
Spitzer et al., 1990). In the years immediately following the
publication of the DSM–III–R, a panel of experts was convened to
outline possible changes to the diagnoses of ODD and CD for the
DSM–IV, which were compiled into the DSM–IV Options Book:
Work in Progress (Task Force on DSM–IV, 1991). The DSM
committee emphasized the need to use more rigorous empirical
information than in previous iterations to justify any changes to the
existing diagnostic system. As part of this process, requests were
made to various investigators to use existing data sets to help
clarify issues regarding the diagnosis of ODD and CD prior to the
official field trials (Frick et al., 1993; Lahey, Loebel, Quay, Frick,
& Grimm, 1992; Loebel, Green, Lahey, Christ, & Frick, 1992; Loebel,
Keenan, Lahey, Green, & Thomas, 1993). These studies provided evidence supporting the DSM–III–R diagnosis of ODD as
an impairing condition that placed children at risk for the develop-
ment of later CD (Cohen, Kasen, Brook, & Stuening, 1991;
Lahey et al., 1993). Evidence also indicated that the two disorders
shared similar etiological risk factors that were simply more pro-
nounced in CD (Lahey et al., 1992), suggesting that ODD and CD
could be viewed as developmental manifestations of the same
underlying disorder, with CD representing the more severe and
later emerging form of psychopathology (Loebel et al., 1993).

Symptom-level analysis of existing data also indicated that elim-
inating the ODD symptom of swearing might help improve the distinc-
tion between ODD and CD, whereas adding more aggressive
symptoms to CD, as well as modifying others, could help
improve its predictive utility (Loebel et al., 1993).

The DSM–IV field trials were initiated with the goal of further
examining the proposed changes that arose from the secondary
analysis of existing data. The DSM–IV field trials sample consisted
of 440 clinic-referred youths who were predominately male
(76.3%) and ranged in age from 4 through 17 years (Frick et al.,
1994). Analyses examining the utility of the old and newly pro-
posed CD symptoms supported the addition of bullying, threaten-
ing, or intimidating others—as well as often staying out after dark
without permission, beginning before 13 years of age—to the
diagnostic criteria for CD. More restricted definitions to existing
CD symptoms were also found to improve positive predictive
power, including specifying that lying needed to be used to obtain
goods or favors from others and that truancy needed to begin
before age 13. The ODD symptom concerning swearing or using
obscene language was eliminated due to mixed support for its
clinical utility in distinguishing between ODD and CD (Frick et
al., 1994; Loebel et al., 1993).

With these modifications in place, analyses were conducted to
determine the optimal diagnostic threshold for the final eight ODD
symptoms and 15 CD symptoms. In contrast to the approach used
in the DSM–III–R field trials, several criteria were used to establish
symptom cutoff points, including associations with measures of
global impairment, test–retest reliability, and convergence with a
clinician’s diagnosis (Lahey et al., 1994). Most of the evidence
pointed toward a dose response relationship between symptom
severity and the level of functional impairment for both ODD and
CD. Therefore, the decision to use specific diagnostic thresholds
was somewhat arbitrary and based largely on minimizing false
positives, improving test–retest reliability, and enhancing concor-
dance with a clinician’s diagnosis. In the end, the diagnostic thresh-
old was set at four of eight symptoms in the past 6 months for
ODD and three of 15 symptoms in the past year for CD, with at
least one CD symptom needing to be present in the last 6 months.
With this new diagnostic criteria, approximately half of the indi-
viduals in the field trial sample who met criteria for CD did not
meet diagnostic threshold for ODD (Lahey et al., 1994), running
counter to previous beliefs that most features of ODD were present
in individuals diagnosed with CD. Despite this finding, the
committee decided to retain the longstanding hierarchical rule prevent-
ing a diagnosis of ODD when CD is present.

The issue of using a subtyping scheme to further delineate
heterogeneity among youths with a CD diagnosis was a matter of
some debate among the DSM–IV committee members (Lahey et
al., 1992). The prior focus on defining subtypes of CD using
aggressive behavior was not embraced due to a lack of objective
criteria for defining the amount and type of aggression required to
meet threshold, as well as concerns about the unreliability of the
subtyping scheme due to fluctuating aggressive symptoms over
time (Lahey et al., 1998). Similar problems arose when attempting
to subtype CD on the basis of the socialized versus unsocialized
distinction.

An alternative approach that did not suffer from these problems
involved using the developmental timing of the individual’s first
CD symptom (Moffitt, 1993). This approach was appealing be-
cause several studies had found that the most severe and persistent
offenders tended to initiate their antisocial behavior in childhood
rather than adolescence, and it eliminated the complexities asso-
ciated with having a dynamic subtyping scheme (Loebel, 1988;
Moffitt, 1993). The DSM–IV field trials data also supported this
distinction (Lahey et al., 1998), with youths with childhood-onset CD (first symptom prior to age 10) having higher levels of aggression and more ODD symptoms than did youths with adolescent-onset CD (first symptom after age 9). This suggested that childhood-onset CD seemed to be capturing many youths who previously fell into the aggressive subtype of CD, solidifying its place as the sole CD subtyping scheme in DSM–IV.

Preparing for the DSM–5

In 2007, the DSM–5 ADHD and Disruptive Behavior Disorders Work Group was formed and began meeting. Members of the work group conducted research reviews to outline pertinent questions that should be addressed to support possible revisions to the DSM–IV criteria for ODD and CD. Although these issues are outlined in further detail elsewhere (Frick & White, 2008; Moffitt et al., 2008), four issues are the primary focus of this current special section.

First, research investigating ODD and CD in girls had been limited in prior clinical and epidemiological longitudinal studies, and the previous DSM field trials had included only a small number of girls. This caused some to question whether the diagnostic criteria for ODD and CD had the same clinical utility in girls as boys and whether gender-specific diagnostic criteria and thresholds should be used (Keenan, Coyne, & Lahey, 2008).

Second, there remains a continuing controversy about the appropriateness of retaining the hierarchical rule that precludes a diagnosis of ODD when the criteria for CD are met, especially because evidence suggests that the overlap between the disorders may not be as substantial as once thought (Lahey et al., 1994). Further, there is emerging evidence that the symptoms of ODD can be clustered into distinct facets (e.g., angry/irritable mood, defiant/headstrong behavior, vindictiveness) that have unique associations with CD and internalizing problems (Stringaris & Goodman, 2009).

Third, a particularly important set of questions being considered for the DSM–5 is how to recognize the developmental associations between CD in children and antisocial personality disorder (ASPD) in adults and, relatedly, how best to incorporate the affective and interpersonal features of psychopathy into these definitions (Hare & Neumann, 2008). In recent years, the use of callous–unemotional (CU) traits (e.g., lack of guilt and empathy) has shown promise as a means of delineating a subtype of CD youths with (a) a particularly severe and recalcitrant form of antisocial behavior and (b) distinct neurological, cognitive, emotional, and social characteristics (Frick & White, 2008). Although this method of subtyping youths with CD appears quite promising, more work is needed to determine how best to incorporate these traits into diagnostic criteria in a valid and useful way.

Fourth, there are continued debates surrounding the relative benefits of a dimensional versus categorical conceptualization of mental disorders in general and ODD and CD specifically. That is, much of the research on ODD and CD uses a dimensional approach to these symptoms (Moffitt et al., 2008), whereas the diagnostic classification emphasizes a categorical approach. The DSM–IV included a severity index for CD, but this has largely been ignored by both researchers and clinicians. In addition, the severity classification for ODD was inexplicably eliminated during the transition from the DSM–III–R to the DSM–IV (American Psychiatric Association, 2000). Thus, an important question for research is to demonstrate what types of dimensional approaches to ODD and CD may enhance the utility of the diagnostic criteria by providing a meaningful metric of severity for both researchers and clinicians.

Building an Evidence Base for the DSM–5

Using six different longitudinal samples, the empirical articles in the current special section were designed to address these questions important for the DSM–5. The use of longitudinal data analysis to establish predictive utility is seen as one of the more important criteria for justifying modifications to the existing diagnostic system (Moffitt et al., 2008). Further, these studies span a large age range and include both epidemiological and clinic samples that are quite diverse in their gender and racial composition.

The first study, by Keenan, Wroblewski, Hipwell, Loeber, and Stouthamer-Loeber (2010), examined the validity of the age of onset distinction and CD symptom threshold in girls and explored whether relational aggression and CU traits have clinical utility in predicting impairment in girls. The study involved a large community sample of girls (N = 2,451) who were assessed annually from age 7 to 15 using multiple informants. Contrary to predictions, adolescent-onset CD was rare in girls, and only about half of those girls diagnosed with CD across all phases had met criteria for ODD at any phase. Findings indicated that the relation between CD symptoms and later impairment was largely linear, suggesting that the existing diagnostic threshold of three symptoms is relatively arbitrary in girls. Although relational aggression and CU traits incrementally added to the prediction of later impairment above and beyond CD symptoms, the best method for incorporating these features into the existing CD criteria was unclear.

Using a birh cohort of 995 New Zealand–born individuals, Fergusson, Boden, and Horwood (2010) focused on comparing dimensional versus categorical conceptualizations of ODD and CD in their ability to predict adult outcomes across multiple domains. The study also examined the unique associations between adolescent ADHD, ODD, and CD symptoms and early adult adjustment and explored whether these relations differed by gender. The results clearly supported the use of a dimensional model, with subclinical levels of ODD and CD representing significant risk factors for a wide variety of negative outcomes. Relatively distinct developmental outcomes in adulthood were also associated with adolescent ODD and CD symptoms. Specifically, ODD symptoms were robustly related to the development of internalizing problems, whereas CD symptoms were a consistent predictor of future antisocial outcomes. Importantly, there was no substantive evidence that these associations varied by gender.

The next study, by Kolko and Pardini (2010), used a clinic sample of 177 children 6–11 years of age diagnosed with either ODD or CD to examine the clinical utility of dimensions of ODD and CU traits for predicting treatment outcomes across a 3-year period. Children who met the proposed DSM–5 threshold for CU traits (Frick & Moffitt, 2010) were more likely to be diagnosed with CD rather than ODD, with the CU subtype of CD representing 59.5% of all childhood-onset CD cases. Although this suggests that the proposed CU subtyping scheme can be used to further subdivide childhood-onset CD cases, children meeting criteria for the CU subtype did not have poorer treatment outcomes. In con-
trast, delineating between different facets of ODD did have significant clinical utility, even after controlling for co-occurring CD. Specifically, the irritability facet of ODD was uniquely associated with posttreatment internalizing problems, whereas the vindictiveness symptom of ODD uniquely predicted posttreatment CD symptoms/diagnosis and delinquent behaviors.

Rowe, Costello, Angold, Copeland, and Maughan (2010) used a longitudinal epidemiological sample of 1,420 boys and girls covering a developmental span from 9 to 21 years of age to examine questions regarding the developmental transitions between ODD and CD, as well as the differential predictive utility of ODD symptom facets. Evidence suggested that the transition from ODD to CD was less common than expected, with nearly half of all children with a CD diagnosis having no prior ODD diagnosis. However, an ODD diagnosis was found to be a risk factor for the development of CD even after controlling for subclinical CD symptoms, particularly for boys. As in other studies in the section, ODD showed stronger predictions to internalizing disorders in early adult life than did CD, and the irritable mood symptoms (e.g., angry or resentful) of ODD were more strongly related to the prediction of anxiety disorders than were the headstrong symptoms (e.g., arguing with adults, defying adults’ requests).

Using three longitudinal data sets, Burke, Waldman, and Lahey (2010) investigated the developmental progression between ODD and CD, the importance of considering subclinical symptoms of ODD and CD, and the predictive utility of features associated with CU traits. Findings indicated that although ODD preceded the development of CD for some youths, there were a substantive number of youths with adolescent-onset CD who had no prior history of ODD. There was also evidence that the use of a modified diagnosis of ODD that combined subclinical symptoms of both ODD and CD identified children with significant global impairment, suggesting that it is important to consider ODD and CD symptoms on a continuum. Lastly, features of interpersonal callousness were found to be a risk factor for developing ASPD by early adulthood in clinic-referred boys, but contrary to expectations, this association emerged only in boys without a CD diagnosis.

The study by McMahon, Witkiewitz, Kotler, and the Conduct Problems Prevention Research Group (2010) also attempted to look at the incremental utility of adolescent CU traits for predicting criminal behavior and antisocial personality in early adulthood using a sample of 754 boys and girls. The findings from this study were much more clear-cut, because adolescent CU traits predicted self-reported delinquency, arrests, and ASPD in early adulthood even after controlling for ODD and CD symptoms and childhood-onset CD. These effects were not moderated by race or gender, further supporting the robustness of these findings. However, the predictive utility of CU traits was examined using a continuous measure rather than by defining a subtype of CD based on CU traits, as proposed for the DSM–V (Frick & Moffitt, 2010).

**Future of ODD and CD in the DSM–5**

The goal of the current special section is not to make any definitive recommendations about changes to the DSM classification system but rather to provide an evidence base that should be considered when updating the criteria for ODD and CD. In the end, there are several factors well beyond the scope of the studies contained here that are important to consider. Decisions will have to be made about the relative weight to give findings from epidemiological versus clinic samples, because study results often do not generalize across these populations, especially when it comes to the overlap between ODD and CD. The relative strengths and weaknesses of these studies will also have to be considered when resolving contradictory results. Tough decisions about defining symptom thresholds and how to handle subclinical symptoms will have to be made in light of the ramifications that these decisions will have in terms of insurance reimbursement and the proliferation of diagnostic labeling. Although any modifications to the diagnosis of ODD and CD in the DSM–5 will be a matter of debate for years to come, as has been the case with all prior DSM revisions, the insights provided by the current special section will help to make the evidence base available to inform these decisions for the DSM–5 far greater than was the case for past DSM iterates.

**References**


doi:10.1097/00004583-199111000-00018


doi:10.1097/00004583-199405000-00011


doi:10.1016/0272-7558(93)90016-F


doi:10.1111/j.1469-7610.2007.01862.x


Received August 12, 2010
Revision received August 17, 2010
Accepted August 17, 2010