Published in final edited form as:

J Abnorm Psychol. 2012 May ; 121(2): 352–359. doi:10.1037/a0024340.

Mind the gap: Assessing impairment among children affected by proposed revisions to the diagnostic criteria for oppositional defiant disorder

Kate Keenan, Ph.D.

Department of Psychiatry and Behavioral Neuroscience, University of Chicago

Abstract

The DSM-5 workgroup for disruptive behavior disorders is considering adopting a frequency threshold for symptoms of Oppositional Defiant Disorder (ODD). In the present study, the impact of substituting the term “*often*” with a specific age-based frequency on impairment and prognosis among preschool children was tested in a longitudinal design. Mutually exclusive groups were created to identify children who met criteria for ODD based on a symptom threshold of “*often*,” as in DSM-4, and those that met criteria for ODD based on a threshold of “*1–2 times per day*,” which approximated the proposal for DSM-5. Comparisons of these groups to each other and to non-diagnosed peers determined the impact of changing the symptom threshold on impairment and prognosis. Close to a third of children who met DSM-4 criteria for ODD did not meet criteria under the alternative diagnosis; African American children were overrepresented in this group. Preschoolers who met DSM-4, but not the alternative criteria had higher rates of ODD, CD, and were more impaired than their non-diagnosed peers at baseline and follow-up. Preschoolers meeting DSM-4 criteria were less impaired than children meeting the alternative ODD criteria at baseline according to parent, but not according to teacher report. No differences could be detected between those meeting DSM-4 and those meeting the alternative criteria in rate of ODD, CD, or impairment at follow-up. Among clinically referred preschool children, changing the symptom threshold for ODD could result in a sizable group of children who would no longer meet diagnostic criteria, despite demonstrating significant morbidity concurrently and prospectively.

Keywords

preschoolers; ODD; diagnosis; longitudinal; DSM-5

Introduction

As one of the most prevalent disorders of childhood and one of the most common reasons for referral of young children to mental health clinics (Lavigne, Lebailly, Hopkins, Gouze, & Binns, 2009; Wilens, Biederman, Brown, Monuteaux, Prince, & Spencer, 2002), Oppositional Defiant Disorder (ODD; American Psychiatric Association, 1994) significantly impacts public health. Childhood ODD is stable and predictive of poor psychiatric outcomes (Cantwell & Baker, 1989). ODD that emerges in the preschool period is also highly stable, with approximately 80% of diagnosed preschoolers continuing to meet criteria for ODD during a 3-year follow-up (Keenan, Danis, Hill et al., 2010). Moreover, ODD is a gateway to many forms of adolescent and adult psychopathology. In the Great Smoky Mountain Study, for example, a diagnosis of ODD during childhood and adolescence predicted young adult depression; depression during childhood and adolescence did not (Copeland, Shanahan,

Costello & Angold, 2009). In the Dunedin Multidisciplinary Health and Development Study, ODD and conduct disorder (CD) were the only childhood disorders that predicted every adult disorder (Kim-Cohen, Caspi, Moffitt, Harrington, Milne, & Poulton, 2003). Consequently, considering changes to the operational definition of ODD must be weighed against the extant data supporting its stability, predictive utility, and capacity for informing prevention efforts for many common mental disorders.

Like most common childhood disorders, including CD, attention deficit hyperactivity disorder, and separation anxiety disorder, ODD is defined by the presence of a specific number of symptoms that persist for a specified period of time; an objective measure of frequency of symptom manifestation is not specified. For DSM-4 ODD (the 4th version of the DSM is denoted by an Arabic numeral to be consistent with DSM-5), 4 symptoms are required to be present during a 6-month period. For each of the 8 symptoms, the frequency is defined as *often*, as in “*often loses temper*.” In fact, nearly all childhood disorders in the DSM-4 lack specific, frequency based definitions of symptom manifestation (elimination and eating disorders are some exceptions).

The DSM-5 child and adolescent disorders workgroup is considering changing the frequency of ODD symptom manifestation to include different sets of frequency criteria that vary by age and by symptom. For children under 5 years of age, behaviors must occur on most days (with the exception of *spiteful and vindictive*). For individuals 5 years or older, the behavior must occur at least once per week (with the exception of *spiteful and vindictive*). At both ages the symptom of spiteful and vindictive should occur at least twice in the past 6 months (American Psychiatric Association, 2010). The proposed changes are designed to increase validity, especially for younger children. There is a concern that a low threshold for frequency or a subjective measure such as “*often*” may result in the misclassification of younger children as disordered. Evidence to date supports the reliability and concurrent and predictive validity of DSM-4 ODD in preschoolers (Keenan, Wakschlag, Danis, Hill, Humphries, Duax, & Donald, 2007; Keenan et al., 2010). Therefore, prior to changing the symptom threshold an empirical comparison of the current and proposed symptom thresholds should be conducted.

There are very few datasets upon which to draw empirically based conclusions regarding operational definitions of symptom thresholds for ODD, and for preschool children in particular. Longitudinal data beginning in the preschool period and extending into school age is optimal for assessing predictive validity. An assessment of frequency for each symptom in addition to the DSM-4 operational definition of “*often*” would also be needed. In the present study, these criteria are met. Preschool children, ages 3 to 5 were followed into school age (ages 6–8). Mothers were asked if children “often” engaged in a specific behavior. Regardless of whether mothers reported that the behavior occurred often, frequency was reported. Thus, diagnoses of ODD based on different symptom thresholds can be compared in terms of concurrent and prospective assessments of ODD and impairment. Specifically, we are interested in the overlap of children meeting DSM-4 and an alternative diagnostic threshold based on frequency of symptom manifestation. We hypothesize that there will be a sizable group of children who will meet DSM-4, but not the alternative criteria, and that all children who meet the alternative criteria will also meet DSM-4 criteria. Once the overlap is specified we further examine the level of impairment and hypothesize that children who meet DSM-4, but not the alternative criteria will show higher levels of impairment compared to children not meeting either criteria. In addition, we expect that children who meet DSM-4, but not the alternative criteria, will be either comparably impaired, or in some comparisons, somewhat less impaired than children meeting the alternative criteria due to the fact that the alternative criteria is based on more frequent manifestations of symptoms.

Methods

Participants

Data on the participants, preschool and follow-up assessments, and the reliability and validity of the diagnostic instrument used in the present study have been published previously (Keenan et al., 2007; 2010). For clarity and ease of interpretation of the data presented in this study, we briefly review those results.

Two hundred twenty-three 3- to 5-year-old children, residing with their biological mothers, were enrolled in the study. Approximately half of the children (n=123) had been referred to a child psychiatry outpatient clinic because of aggression, defiance, and/or problems controlling temper. The other half (n=100) was recruited through the general pediatric clinic, which had similar demographic composition. Family income was set at a cut-off of at or below 250% of the U.S. poverty level by family size in order to recruit a continuum of families living in low-income environments.

Of the 152 consecutive clinic referrals 130 (85.5%) agreed to participate, 8 (5.3%) refused to participate, 7 (4.6%) were unable to be scheduled for a visit, 5 (3.3%) sought services elsewhere, and 2 (1.3%) never completed the recruitment process. Seven children who were recruited into the study were later excluded from the analytic sample because of serious medical problems that had onsets prior to their referral (i.e., brain tumor, genetic syndrome, and seizure disorder).

Caregivers of non-referred children responded to fliers in the waiting rooms of pediatric clinics and were screened to establish that the caregiver was not seeking an evaluation for disruptive behavior problems. Of the 109 families who responded and were eligible for the study based on demographic characteristics that were consistent with the clinic sample (i.e., race, age, and sex), 100 (91.7%) agreed to participate, 1 (0.9%) refused, and 8 (7.3%) were unable to be scheduled. Eighty-two percent of all participating children were African American, and 52.5% were male. The age distribution was as follows: 79 (35.4%) three-year olds, 75 (33.6%) four-year olds, and 69 (30.9%) five-year olds. The Institutional Review Board approved all study procedures. Written informed consent was obtained after providing a complete description of the study to the subjects.

Children were followed annually for a period of three years, up through school age. Multiple attempts at contacting subjects, home visits for families without phones, and a willingness on the part of the staff to reschedule yielded high retention at each follow-up: 92.4% at 12 months (ages 4–6), 93.3% at 24 months (ages 5–7), and 91.5% at 36 months (ages 6–8). There were no significant differences between those participants who were and were not seen in each wave in terms of race, sex, referral status, or DSM-4 ODD and CD diagnosis at baseline. The biological mother was the informant in nearly all cases over the course of the study. In the final wave, only 3 (1.4%) of the 204 subjects retained had informants other than the biological mother.

Disruptive behavior symptoms and diagnoses

Bachelor's level research assistants, who were blind to the referral status of the child, administered the Kiddie-Disruptive Behavior Disorder Schedule (K-DBDS; Keenan et al., 2007) to caregivers. Different research assistants administered the preschool and school age assessments. The K-DBDS was developed to assess disruptive behavior disorders in young children using developmentally appropriate operational definitions of symptoms. DSM-4 ODD symptoms were unchanged from how they are described in DSM-4. Of the 15 DSM-4 CD symptoms, four were not assessed because of the lack of validity: (1) breaking into a house, car, or building; (2) running away from home overnight; (3) often staying out late;

and (4) truancy. The wording was modified for six symptoms. For example, “*Has stolen items of nontrivial value without confronting a victim*” was modified to read, “*Has stolen from a store or purse on more than one occasion.*” “*Has used a weapon that can cause serious physical harm to others*” was modified to read, “*Has used an object (such as a stick, rock, or knife) to hurt or to try to hurt someone.*” A complete description of modifications and examples of symptom manifestation has been published (Keenan et al., 2007).

The one-week test-re-test reliability of the K-DBDS was conducted in an independent sample of 31 referred and non-referred preschoolers. Test-retest reliability for total number of ODD and CD symptoms was high ($ICC > .75$). Inter-rater reliability for the total number of symptoms was high ($ICC > .95$), as was the inter-rater reliability for diagnoses ($kappa > .90$). Reliability was not affected by age or sex of the child. Validity of preschool diagnoses was demonstrated via associations with impairment, such as parental ratings of global impairment, and differentiation between referred and non-referred children (Keenan et al., 2007). A licensed clinical psychologist reviewed the K-DBD and impairment ratings with the research team for each subject.

DSM-4 and DSM-Alternative (DSM-A) Diagnostic Groups

For all ODD symptoms the mother was first asked whether the child *often* engaged in a behavior (e.g., “does she often lose her temper”). If a mother responded negatively to this question she was asked if the child *ever* engaged in the behavior. All mothers who reported that children *often* or *ever* engaged in a behavior were asked frequency: *rarely*, *a few times per month*, *a few times per week*, *1–2 times per day*, or *many times per day*. For the present study, DSM-4 ODD is defined by 4 symptoms endorsed at the level of *often* during a period of 6 months. An alternative symptom threshold (DSM-A), which approximates the proposed DSM-5 changes, was defined by endorsements at the level of *1–2 times per day* for all symptoms with the exception of spiteful or vindictive, for which an endorsement at the level of *rarely* during the past 6 months was deemed sufficient.¹ ODD and CD diagnosis at follow-up was based on DSM-4 criteria.

Functional Impairment

Caregivers completed the non-clinician version of the Children’s-Global Assessment Scale (C-GAS; Setterberg, Bird, Gould, Shaffer, & Fisher, 1992), which has been validated for use with parents (Bird, Andrews, Schwab-Stone, et al., 1996). Scores on the C-GAS range from 1 to 100 with each decile containing a description of whether and how much the child’s behavior negatively impacts functioning, with lower scores reflecting greater levels of impairment. A score at or below 60 was used to indicate clinically significant impairment (Bird, 1999).

School data

The Early Childhood Inventory (ECI-4; Sprafkin & Gadow, 1996) a behavior rating scale that screens for DSM-4 emotional and behavioral disorders in children was used as the teacher report measure. Items are rated as *never*, *sometimes*, *often*, or *very often*. The ECI has good reliability and predictive and concurrent validity (Sprafkin, Volpe & Gadow, 2002). For a diagnosis of ODD, symptoms were counted as positive if endorsed at the level of ‘often’ or ‘very often.’ The absence of time-based frequency data precluded an assessment of alternative symptom thresholds by teacher report. Data from the ECI were successfully collected from 92.4% of the participants at baseline and from 83.0% and 76.2%

¹Note that in the proposed revision to DSM-5 ODD, a different frequency threshold is suggested for 3–4 year olds versus 5-year olds. In the present sample, there was no evidence of an interaction between age and frequency on impairment for any of the ODD symptoms. Thus, the same frequency threshold was used for all preschoolers for diagnosing DSM-A.

of participants at the 12- and 24-month follow-ups, respectively. ECI data were only collected from a minority of participants at the 36-month follow-up due to funding constraints; follow-up data are only presented for the 12- and 24-month assessments for ODD diagnoses on the ECI, through ages 5 to 7 years.

Mothers reported on whether the child had been expelled from preschool. At baseline, 11 children had been expelled, all of who were in the clinically referred group (8.9% of the clinically referred group). Only 7 children were expelled during follow-up, resulting in cell sizes too small to analyze. Thus, expulsion from preschool is used for analysis only at baseline.

Results

Distribution across diagnostic groups

Among the 223 preschoolers, 89 (39.9%) met DSM-4 criteria for ODD (91.0% of who were from the clinically referred group). There were no significant sex or race differences between those meeting and those not meeting DSM-4 criteria for ODD. Among those meeting DSM-4 criteria, the most commonly endorsed symptoms of ODD were defiance (100.0%), losing temper (94.4%), annoying others on purpose (84.3%), and blaming others for her own mistakes (78.7%). Using the DSM-Alternative (DSM-A) criteria for symptom threshold as described above, 62 preschoolers (27.8%) met criteria for ODD, a reduction in rate of approximately 33%. As shown in Figure 1, the decrease in the number of children meeting criteria is due to the fact that a frequency of “*a few times per week*” was a common response for many symptoms that were reported to have occurred *often*. Among those meeting DSM-A criteria, the most commonly endorsed symptoms of ODD were defiance (98.4%), annoying others on purpose (86.6%), and blaming others for her own mistakes (85.5%), and arguing with adults (84.9%).

Mutually exclusive groups were created to identify preschool children who met criteria based on DSM-A ODD ($n = 63$) compared to those meeting criteria based on DSM-4 but not DSM-A ($n = 28$). Comparisons of these groups to each other and to non-diagnosed peers ($n = 132$) provide an opportunity to determine the impact of changing the criteria for symptom threshold on the 28 cases that would no longer meet criteria for ODD. As shown in Table 1, nearly all (60/62) participants who met criteria for DSM-A ODD also met criteria for DSM-4 ODD (2 preschoolers had reported frequencies of *1–2 times per day*, but were not reported to *often* engage in the behaviors). The primary focus is on the 29 children who met DSM-4 criteria, but not the DSM-alternative criteria given the primary aim: to determine whether children in the DSM-4 only group differ from non-diagnosed peers on clinically significant indicators. If they do differ, then the fact that these children would no longer meet criteria would be a concern. In addition, although we would expect that the children in the DSM-A ODD group would demonstrate greater impairment than those in the DSM-4 only group given the selection on frequency of symptom manifestation, if the magnitude of difference between the two groups is relatively small or non-significant, this would increase the level of concern for the children falling into the gap.

Demographic factors

Demographic factors were examined first including age, sex, and race. Comparisons using chi-square analysis revealed that age ($\chi^2 [4] = 3.75, p > .10$) and sex ($\chi^2 [2] = 1.66, p > .10$) were not associated with diagnostic groups. Distribution of African American children was associated with diagnostic group ($\chi^2 [2] = 7.45, p < .05$): 85.6% of non-diagnosed, 89.7% of DSM-4 only, and 71.0% of DSM-A were African American children.

Comparison of DSM-4 and DSM-A ODD on concurrent measures

Separate multinomial logistic regressions were conducted to compare children in the DSM-4 only group to non-diagnosed peers and to children meeting DSM-A criteria on parent-rated C-GAS at or below 60, diagnosis of DSM-4 CD by parent, diagnosis of DSM-4 ODD by teacher, and expulsion from preschool. The multinomial logistic regression is appropriate for unordered categorical outcomes greater than two, and provides a parsimonious approach to conducting comparisons of one group, in this case children in the DSM-4 only group, to the other two groups, thus reducing the risk of a type I error.

Compared to non-diagnosed peers, preschoolers in the DSM-4 only group were more than 10 times more likely to have a C-GAS at or below 60 ($OR = 10.9 [3.7 - 31.8]$, $p < .001$), and were more than 8 times more likely to meet criteria for CD ($OR = 8.5 [3.4 - 21.6]$, $p < .001$). There was no significant increase in the risk of meeting criteria for ODD by teacher report, but the risk of being expelled from preschool was significantly higher in the DSM-4 only group compared to non-diagnosed peers ($OR = 9.4 [2.09 - 49.65]$, $p < .001$) (Table 2).

Preschoolers meeting the proposed DSM-A criteria were more likely than preschoolers in the DSM-4 only group to meet criteria for CD ($OR = 4.0 [1.6 - 10.5]$, $p < .01$), and were more likely to have a C-GAS at or below 60 ($OR = 2.6 [1.04 - 6.42]$, $p < .05$). There was no difference between preschoolers meeting the proposed DSM-A criteria and those in the DSM-4 only group in rate of ODD diagnosis based on teacher report or on likelihood of expulsion from preschool (Table 2).

Comparison of DSM-4 and DSM-A ODD on prospective measures

Diagnosis at follow-up was determined by meeting DSM-4 criteria for ODD or CD by parent report in at least 1 of the three annual follow-up assessments (through ages 6–8 years). For teacher reported ODD, follow-up status was based on meeting criteria for DSM-4 ODD in at least one of the two annual assessments (through ages 5 to 7 years). Compared to children who did not meet criteria for ODD at preschool age, children who met criteria for DSM-4 ODD as preschoolers were significantly more likely to meet criteria for ODD ($OR = 6.3 [2.3 - 15.3]$, $p < .001$) and CD ($OR = 5.7 [2.1 - 15.5]$, $p < .001$) by parent report, and ODD by teacher report ($OR = 3.7 [1.3 - 10.6]$, $p < .05$) during follow-up. No statistically significant differences in diagnosis at follow up between children who met DSM-4 criteria as preschoolers and those who met DSM-A criteria as preschoolers could be detected (see Table 3).

Clinically significant impairment, as defined by a C-GAS at or below 60, was examined as a function of preschool diagnostic group in each of the three follow up assessments. As shown in Table 4, children who met criteria for DSM-4 ODD were more likely to be impaired at the 12- ($OR = 3.4 [1.11 - 10.60]$, $p < .05$), 24- ($OR = 4.1 [1.15 - 14.40]$, $p < .05$), and 36-month follow-ups ($OR = 5.3 [1.41 - 19.75]$, $p < .05$) than their non-diagnosed peers. No statistically significant differences in impairment at follow-up between preschoolers who met DSM-4 and preschoolers who met DSM-A criteria for ODD could be detected.

Discussion

Refinements to diagnostic criteria are warranted when there is evidence to support misclassification of individuals. In the case of preschool children, for whom diagnostic validity for the disruptive behavior disorders has only recently accumulated, the consideration is significant. Although there was initial resistance to the application of diagnostic constructs to the behavioral and emotional functioning of preschoolers (McClellan & Speltz, 2003; Thomas & Guskin, 2001) there is now a solid base of evidence supporting the hypothesis that psychopathology can present early in life and, that when

tested, the diagnostic constructs described in the DSM such as depression (Luby, Si, Belden, Tandon, & Spitznagel, 2009), autism (Lord, Risi, DiLavore, Shulman, Thurm, & Pickles, 2006), and ADHD (Lahey, Pelham, Loney, et al., 2004), are reliable and valid for young children. It is reasonable, however, to raise the question of whether developmental modifications will improve the reliability and validity of a diagnosis. With few exceptions, the assessment of childhood disorders in the DSM does not take age into consideration. For disorders that include behaviors that are not indicative of pathology *per se*, as is the case for ODD, CD, ADHD, and separation anxiety efforts to ensure that the phenotype is clinically meaningful across development are needed, especially in the context of a decrease in the age of referral for mental health services (Howell & Teich, 2008). The rationale for changing the ODD symptom threshold in DSM-5 is that objective criteria would provide clinicians with guidance on the level at which behaviors, which occur to some degree in non-disordered individuals, are consistent with psychopathology (American Psychiatric Association, 2010). The workgroup also recognized the potential disadvantages of developing objective frequency criteria including the possibility that objective thresholds that generalize across age, gender, and cultural groups may not exist (American Psychiatric Association, 2010).

The aims of the present study were relatively narrow but do bear on the question raised by the proposed changes for DSM-5 ODD with regard to concerns about misclassifying younger children as disordered. Specifically, the aims were to determine how many preschool children would no longer meet criteria for ODD when an alternative symptom threshold that approximated the proposed DSM-5 changes was applied, and to compare that group to children without a disorder and to children meeting the alternative criteria in terms of clinical indicators.

In examining the overlap between DSM-4 and the alternative DSM-A, nearly all (61/63) participants who met criteria for DSM-A ODD met criteria for DSM-4 ODD (2 children had reported frequencies of *1–2 times per day*, but were not reported to *often* engage in the behaviors). Thus, there do not appear to be concerns that the DSM-4 criteria are too stringent, or that preschool children are being under-identified. Approximately one third of preschoolers who met criteria for DSM-4 ODD did not meet the alternative ODD criteria. This is a substantial minority of children, and the question was raised as to whether the DSM-A criteria were too stringent leading to preschoolers in need of services being under-identified. Comparisons at baseline revealed that preschoolers who met DSM-A criteria were more impaired and were more likely to meet criteria for CD than preschoolers who met DSM-4, but not DSM-A criteria. These data suggest that changing the symptom threshold results in identifying a more impaired group of children. This is not surprising given that the groups were defined by severity of frequency of symptom manifestation.

The next question is whether these children are in fact misclassified via the DSM-4 system. One approach is to compare them to non-diagnosed peers. In this sample, children who met DSM-4 criteria were more than 10 times more likely to be impaired, more than 8 times more likely to be diagnosed with CD, and more than 9 times more likely to be expelled from preschool than their non-diagnosed peers, suggesting that they are not misclassified. Another approach is to use prospectively collected data to assess long-term outcome. This is an approach that may only be available through the use of extant data and not available in the context of the DSM-5 field trials. Based on 3 years of annual assessments from mothers and 2 years from teachers, comparisons between diagnostic groups revealed significant differences between the children who met DSM-4, but not DSM-A criteria as preschoolers and non-diagnosed peers in parent and teacher reported diagnoses of ODD, parent diagnosis of CD, and parent-reported impairment. No significant differences were could be detected between children who met DSM-4 and children who met DSM-A criteria for ODD on any of the follow-up measures.

By definition, children who would be left in the ‘gap’ if the symptom threshold moved from *often* to *daily* or *most days* comprise the less severe end of the spectrum of those who met DSM-4 criteria. Children in the entire DSM-4 group demonstrated a range of symptom frequency from *rarely* to *many times per day* as shown in Figure 1, and nearly all of the children who met DSM-A also met criteria for DSM-4. The ‘gap’ group was comprised of a subset of the children meeting DSM-4 criteria whose symptoms manifestations ranged from *rarely* to *a few times per week*. Despite the decreased variability in reported frequency of symptom manifestation, these children displayed significant levels of disorder and impairment currently and over time. The term *often* appeared to function quite well in identifying preschool children who were more impaired than their non-diagnosed peers. In addition, although the term *often* may obscure differences in severity at baseline, such differences did not affect prognosis at school age in terms of diagnosis of ODD, CD, or impaired functioning.

Notably, the loss of cases was greater among African American children than among European American children. This race effect could be due to cultural differences that lead to different operational definitions of “often.” Alternatively, race, which even in the present study which is oversampled for low-income is confounded with SES, may reflect the impact of having fewer resources and increased work demands outside of the home, which lead to less reliable reports of specific behavioral frequencies than for global reports (e.g., *often*). In the context of research highlighting unmet need for mental health services for African American children living in urban poverty (Elster, Jarosik, VanGeest, et al., 2003; Garland, Lau, Yeh, et al., 2005), the possibility that a change in the operational definition of a diagnosis could lead to fewer African American children meeting criteria for a disorder, despite seeking services, is a considerable risk that needs to be weighed against potential benefits. In summary, changing the symptom frequency from *often* to *most days*, as proposed for DSM-5 could result in a substantial number of young children not being eligible for mental health services despite significant impairment and continued risk for psychopathology, and a disproportionate number of those falling into the gap would be African American children.

A limitation of the present study is that the answer choices on the diagnostic instrument used did not map exactly onto the proposed criteria for DSM-5 ODD: *1–2 times per day* was used instead of the proposed *most days* for all symptoms except for spiteful and vindictive, for which *rarely in the past 6 months* was used instead of *at least twice within the past months*. It is possible that such differences led to findings that inflated the impact of changing the symptom threshold on clinical indicators. That being said, a comparison of DSM-4 ODD and the proposed DSM-5 ODD using the exact wording would be useful data to have in resolving this issue. In fact, prior to adopting a symptom threshold, a comprehensive examination of optimal cut-points for each symptom should be undertaken. In the present study, the modal frequency of symptom manifestation among clinically referred preschoolers ranged from *1–2 times per week* for losing temper, spiteful/vindictive and angry to *many times per day* for defiance, blaming others, and arguing with adults. A specific frequency based symptom threshold, therefore, may need to be set for each symptom.

Although there was a sizable group of children with DSM-4 ODD ($n = 89$, 39.9%), when creating mutually exclusive subgroups the groups were relatively small, with 29 subjects in the DSM-4 subgroup, and 62 subjects in the DSM-A subgroup. This may have affected our ability to detect statistically significant differences for some comparisons. All but one of the comparisons between DSM-4 ($n = 29$) and non-diagnosed peers were statistically significant. Thus, for the primary aim of the study, to determine whether children falling into

a gap as a result of a change in symptom threshold differed significantly from non-diagnosed peers, the analyses appear to have been sufficiently powered.

Finally, we recognize that a limitation of the present study is that the sample was comprised of clinically referred children and healthy controls. The vast majority of cases of ODD were among children who were clinically referred, and the number of cases in the non-referred sample did not provide sufficient power to test the effects of changing the symptom threshold on non-referred children. There is a possibility that the results would look different in a non-clinical sample from the results generated in this population. However, this is a population for whom establishing the validity of these disorders is critical: above all other considerations, a clinical disorder must be valid for children seeking mental health services.

As work on DSM-5 progresses, the results of the present study suggest that one needs to be mindful of the possibility that changes in phenotype will have consequences for younger children and for minority youth. By increasing the pathways to specialized mental health services, children who manifest symptoms early in life may have an opportunity to take advantage of efficacious interventions prior to the accumulation of multiple deficits that result from living with a mental disorder. Similarly, there may be a possibility of reducing the overrepresentation of African American youth in special education and juvenile justice systems by recognizing and treating mental disorders earlier in development. Changes to DSM-4 ODD might disproportionately affect these groups of children, and may create barriers to accessing services for such vulnerable populations. Clearly, modifications to diagnostic nosology that raise the threshold for severity, either by increasing the frequency of symptom manifestation or raising the symptom threshold, can result in a more impaired, but smaller group of children. The question for the DSM-5 work groups is whether such modifications will increase the potential to improve mental health outcomes for children and facilitate research on causal factors and preventive interventions.

Acknowledgments

This project was supported by NIMH grant R01 MH62437. The author would like to thank the families and staff of the Chicago Preschool Project, Dr. Paul Frick of the University of New Orleans and the members of the DSM-5 ADHD and Disruptive Behavior Disorders Work Group.

References

- American Psychiatric Association. [Accessed October, 2010] DSM-5 Development. Oppositional defiant Disorder. <http://www.dsm5.org/>
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4. Washington, D.C: American Psychiatric Press; 1994.
- Bird HR, Andrews H, Schwab-Stone M, Goodman S, Dulcan M, Richters J, Rubio-Stipec M, Moore RE, Chiang P, Hoven C, Canino G, Fisher P, Gould MS. Global measures of impairment for epidemiologic and clinical use with children and adolescents. *International Journal of Methods in Psychiatric Research*. 1996; 6:295–307.
- Bird, HR. The assessment of functional impairment. In: Shaffer, D.; Lucas, C.; Richters, JE., editors. *Diagnostic Assessment in Child and Adolescent Psychopathology*. New York: Guilford Press; 1999. p. 209-229.
- Cantwell DP, Baker L. Stability and natural history of DSM-III childhood diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1989; 28:691–700. [PubMed: 2793796]
- Copeland WE, Shanahan L, Costello EJ, Angold A. Childhood and adolescent psychiatric disorders as predictors of young adult disorders. *Archives of General Psychiatry*. 2009; 66:764–72. [PubMed: 19581568]

- Elster A, Jarosik J, VanGeest J, Fleming M. Racial and ethnic disparities in health care for adolescents: A systematic review of the literature. *Archives of Pediatric and Adolescent Medicine*. 2003; 157:867–874.
- Garland AF, Lau AS, Yeh M, McCabe KM, Hough RL, Landsverk JA. Racial and ethnic differences in utilization of mental health services among high-risk youths. *American Journal of Psychiatry*. 2005; 162:1336–1343. [PubMed: 15994717]
- Howell EM, Teich J. Variations in Medicaid mental health service use and cost for children. *Administration and Policy in Mental Health and Mental Health Services Research*. 2008; 35:220–228. [PubMed: 18259853]
- Keenan K, Danis B, Hill C, Humphries M, Boeldt D, Chen D, Coyne C, Davis S, Donald R, Duax J, Hart K, Kampani S, Perrot J, Strickland J. Evidence for the predictive validity of DSM-IV oppositional defiant and conduct disorders diagnosed in preschoolers. *Journal of Child Psychology and Psychiatry*. 2010; 52:47–55. [PubMed: 20738448]
- Keenan K, Wakschlag LS, Danis B, Hill C, Humphries M, Duax J, Donald R. 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*. 2007; 46:457–468. [PubMed: 17420680]
- Kim-Cohen J, Caspi A, Moffitt TE, Harrington H, Milne BJ, Poulton R. Prior juvenile diagnoses in adults with mental disorder: Developmental follow-back of a prospective-longitudinal cohort. *Archives of General Psychiatry*. 2003; 60:709–17. [PubMed: 12860775]
- Lahey BB, Pelham WE, Loney J, Kipp H, Ehrhardt A, Lee SS, Willcutt EG, Hartung CM, Chronis A, Massetti G. Three-year predictive validity of DSM-IV attention deficit hyperactivity disorder in children diagnosed at 4–6 years of age. *American Journal of Psychiatry*. 2004; 161:2014–2020. [PubMed: 15514401]
- Lavigne JV, LeBailly SA, Hopkins J, Gouze KR, Binns HJ. The prevalence of ADHD, ODD, depression, and anxiety in a community sample of 4-year-olds. *Journal of Clinical Child and Adolescent Psychology*. 2009; 38:315–28. [PubMed: 19437293]
- Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A. Autism from 2 to 9 years of age. *Archives of General Psychiatry*. 2006; 63:694–701. [PubMed: 16754843]
- Luby JL, Si X, Belden AC, Tandon M, Spitznagel E. Preschool depression: Homotypic continuity and course over 24 months. *Archives of General Psychiatry*. 2009; 66:897–905. [PubMed: 19652129]
- McClellan J, Speltz M. Psychiatric diagnosis in preschool children [Letter to the editor]. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2003; 42:129.
- Setterberg, S.; Bird, H.; Gould, M.; Shaffer, D.; Fisher, P. Parent and interviewer versions of the Children's Global Assessment Scale. New York: Columbia University; 1992.
- Sprafkin J, Volpe RJ, Gadow KD, Nolan EE, Kelly K. A DSM-IV-referenced screening instrument for preschool children: The Early Childhood Inventory-4. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2002; 41:604–612. [PubMed: 12014793]
- Sprafkin, J.; Gadow, KD. Early Childhood Symptom Inventories Manual. Stony Brook, New York: Checkmate Plus; 1996.
- Thomas JM, Guskin KA. Disruptive behavior in young children: What does it mean? *Journal of the American Academy of Child and Adolescent Psychiatry*. 2001; 40:44–5116. [PubMed: 11195561]
- Wilens TE, Biederman J, Brown S, Monuteaux M, Prince J, Spencer TJ. Patterns of psychopathology and dysfunction in clinically referred preschoolers. *Journal of Developmental and Behavioral Pediatrics*. 2002; 23:S31–S36. [PubMed: 11875288]

Figure 1.
Frequency distribution for ODD symptoms reported to occur “often” among 3–5 year old children

Table 1

Distribution of sample across diagnostic groups (n = 223)

DSM-4 ODD	DSM-A ODD	
	Absent	Present
Absent	132 (59.2%)	2 (0.9%)
Present	29 (12.6%)	60 (27.4)

Note: DSM-4 ODD = 4 symptoms endorsed at the level of “often” during a period of 6 months; DSM-A ODD = 4 symptoms endorsed at a minimum of 1–2 times per day, with the exception of spiteful/vindictive, which required a minimum endorsement of rarely.

Table 2

Concurrent impairment and diagnosis as a function of preschool diagnostic groups

	Preschool diagnostic groups			Comparisons	
	No ODD (n = 132)	DSM-4 (n = 29)	DSM-A (n = 62)	No ODD versus DSM-4	DSM-4 versus DSM-A
Concurrent impairment and diagnosis					
C-GAS 60 (parent)	N (%) 7 (5.3)	N (%) 11 (37.9)	N (%) 38 (61.3)	OR (95% CI) 10.9 (3.75 – 31.77) ***	OR (95% CI) 2.6 (1.04 – 6.42) *
DSM-4 CD (parent)	13 (9.8)	14 (48.3)	49 (79.0)	8.5 (3.38– 21.57) ***	4.0 (1.56 – 10.45) **
DSM-4 ODD (teacher) ^a	11 (8.7)	6 (20.7)	15 (29.4)	2.7 (0.92 – 8.12)	1.6 (0.54 – 4.71)
Expelled from preschool ^b	0 (0.0)	3 (10.3)	8 (13.6)	9.4 (2.09 – 49.65) ***	1.3 (0.33 – 5.56)

Note: ODD = oppositional defiant disorder; CD = conduct disorder; C-GAS = Children's - Global Assessment Scale.

^aN for teacher data: No ODD = 126; DSM-4 = 29; DSM-A = 51;

^bBecause confidence intervals cannot be calculated when one cell contains no cases, the 95%CI around the odds ratio for the comparison between No ODD versus DSM-4 is a conservative estimate based on assigning a positive response to two randomly selected cases from the No ODD group. By multinomial logistic regression:

*** p < .001;
** p < .01;
* p < .05

Table 3
Oppositional Defiant and Conduct Disorders at follow-up as a function of preschool diagnostic groups

	Preschool diagnostic groups			Comparisons	
	No ODD (n = 119)	DSM-4 (n = 28)	DSM-A (n = 58)	No ODD versus DSM-4	DSM-4 versus DSM-A
Diagnosis at Follow-up	N (%)	N (%)	N (%)	OR (95% CI)	OR (95% CI)
DSM-4 ODD (parent) ^a	30 (25.2)	19 (67.9)	50 (86.2)	6.3 (2.56 – 15.32) ***	2.9 (1.00 – 8.80)
DSM-4 CD (parent) ^a	11 (9.9)	10 (38.5)	34 (61.8)	5.7 (2.08 – 15.54) ***	2.6 (0.99 – 6.76)
DSM-4 ODD (teacher) ^b	13 (13.3)	8 (36.4)	14 (38.9)	3.7 (1.31 – 10.64) *	1.1 (0.37 – 3.34)

^aDiagnoses based on parent-report assessed over 36 months;
^bDiagnoses based on teacher-report assessed over 24 months; By multinomial logistic regression:
*** p < .001;
* p < .05.

Table 4
Parent C-GAS < 60 at 12-, 24-, and 36-month follow-up as a function of preschool diagnostic group

	Preschool diagnostic groups			Comparisons	
	No ODD (n = 119)	DSM-4 (n = 28)	DSM-A (n = 58)	No ODD versus DSM-4	DSM-4 versus DSM-A
Impairment at Follow-up	N (%)	N (%)	N (%)	OR (95% CI)	OR (95% CI)
12 months	9 (7.4)	6 (21.4)	20 (35.7)	3.4 (1.11 – 10.60) *	2.0 (0.71 – 0.90)
24 months	6 (4.6)	5 (17.2)	18 (32.1)	4.1 (1.15 – 14.40) *	2.3 (0.75 – 6.93)
36 months	5 (4.1)	5 (18.5)	19 (33.9)	5.3 (1.41 – 19.75) *	2.3 (0.74 – 6.91)

By multinomial logistic regression:
* p < .05