

Is ADHD Diagnosed in Accord With Diagnostic Criteria? Overdiagnosis and Influence of Client Gender on Diagnosis

Katrin Bruchmüller
University of Basel

Jürgen Margraf and Silvia Schneider
University of Bochum

Objective: Unresolved questions exist concerning diagnosis of ADHD. First, some studies suggest a potential overdiagnosis. Second, compared with the male–female ratio in the general population (3:1), many more boys receive ADHD treatment compared with girls (6–9:1). We hypothesized that this occurs because therapists do not adhere to *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV*) and International Classification of Diseases (10th rev.; ICD–10) criteria. Instead, we hypothesized that, in accordance with the representativeness heuristic, therapists might diagnose attention-deficit/hyperactivity disorder (ADHD) if a patient resembles their concept of a prototypical ADHD child, leading therapists to overlook certain exclusion criteria. This may result in overdiagnosis. Furthermore, as ADHD is more frequent in males, a boy might be seen as a more prototypical ADHD child and might therefore receive an ADHD diagnosis more readily than a girl would. **Method:** We sent a case vignette to 1,000 child psychologists, psychiatrists, and social workers and asked them to give a diagnosis. Four versions of the vignette existed: Vignette 1 (ADHD) fulfilled all *DSM–IV/ICD–10* criteria of ADHD. Vignettes 2–4 (non-ADHD) included several ADHD symptoms but stated other ADHD criteria were nonfulfilled. Therefore, an ADHD diagnosis could not be given. Furthermore, boy and girl versions of each vignette were created. **Results:** In Vignettes 2–4 (non-ADHD), 16.7% of therapists diagnosed ADHD. In the boy version of these vignettes, therapists diagnosed ADHD around 2 times more than they did with the girl vignettes. **Conclusions:** Therapists do not adhere strictly to diagnostic manuals. Our study suggests that overdiagnosis of ADHD occurs in clinical routine and that the patient’s gender influences diagnosis considerably. Thorough diagnostic training might help therapists to avoid these biases.

Keywords: ADHD, overdiagnosis, bias, gender, diagnosis
Supplemental materials: <http://dx.doi.org/10.1037/a0026582.supp>

Attention-deficit/hyperactivity disorder (ADHD) affects a broad range of society with a prevalence rate between 2% and 7% (American Psychiatric Association [APA], 1994; August, Realmuto, MacDonald, Nugent, & Crosby, 1996; Esser, Schmidt, & Woerner, 1990; Fergusson, Horwood, & Lynskey, 1993; Kashani, Orvaschel, Rosenberg, & Reid, 1989). Several therapeutic interventions have been developed and have been empirically shown to be effective in treating ADHD (Purdie, Hattie, & Carroll, 2002; Swanson et al., 2008a, 2008b). If patients are to benefit from such treatments, a thorough and reliable diagnosis is an important precondition.

A scientific and public debate is ongoing regarding the question of whether ADHD is overdiagnosed in children (e.g., “Parenting With Pills,” 2004; Sciotto & Eisenberg, 2007). Only a few empir-

ical studies have addressed this important question, and they have found a trend for potential overdiagnosis. For example, Cotuono (1993) reexamined 92 children previously referred to a specialized ADHD clinic. After comprehensive evaluations, only 22% of the sample was given a primary diagnosis of ADHD and only 37% of the sample was given a secondary diagnosis of ADHD. Desgranges, Desgranges, and Karsky (1995) found that among clinic referrals for suspected ADHD, 62% were not confirmed as ADHD cases after further diagnostic evaluation. A study of Wolraich et al. (1990) showed similar results. These studies suggest that ADHD might, in some cases, be overdiagnosed in clinical routine. However, as it is not clear whether all of the children in these studies had previously been diagnosed with ADHD, the data do not provide sufficient evidence of overdiagnosis of ADHD. To definitely answer the question of ADHD overdiagnosis, one needs to systematically investigate whether professionals responsible for the determination of ADHD diagnosis, diagnose ADHD in cases where diagnostic criteria for ADHD are not fulfilled. Unfortunately, this type of study has not been conducted so far (Sciotto & Eisenberg, 2007).

A second unresolved question regarding the diagnosis of ADHD concerns the male-to-female ratio (Gaub & Carlson, 1997; Gershon, 2002). In representative population-based studies, the male-to-female ratio of ADHD is approximately 3:1 (Barkley, 2006; Gaub & Carlson, 1997), whereas in clinical samples, the male-to-

This article was published Online First December 26, 2011.

Katrin Bruchmüller, Department of Clinical Psychology and Psychiatry, University of Basel, Basel, Switzerland; Jürgen Margraf, Department of Clinical Psychology and Psychotherapy, University of Bochum, Bochum, Germany; Silvia Schneider, Department of Clinical Child and Youth Psychology, University of Bochum.

Correspondence concerning this article should be addressed to Katrin Bruchmüller, Department of Clinical Psychology and Psychiatry, University of Basel, Missionsstrasse 60–64, CH-4055 Basel, Switzerland. E-mail: katrin.bruchmueller@unibas.ch

female ratio is mostly between 5:1 and 9:1 (APA, 1994; Gaub & Carlson, 1997; Gershon, 2002; Goodyear & Hynd, 1992; Lahey et al., 1994; Sandberg, 2002). These numbers suggest that far more boys than girls receive treatment for ADHD compared with the ratio found in population-based samples. This remarkable difference is mostly explained by differences in the expression of the disorder among boys and girls. Several studies suggest that girls with ADHD show fewer behavioral and conduct problems and are less impulsive than boys with ADHD (Berry, Shaywitz, & Shaywitz, 1985; Hartung, Milich, Lynam, & Martin, 2002; Newcorn et al., 2001). In a meta-analysis on gender differences in ADHD, Gaub and Carlson (1997) concluded that girls with ADHD were less hyperactive and showed less aggressive behavior than boys with ADHD. Gershon (2002) found similar results in his more current meta-analysis. Furthermore, girls are more likely than boys to be the predominantly inattentive subtype of ADHD (Biederman et al., 2002). The symptoms of inattention might be less likely to be disruptive in the classroom situation. This might lead to less frequent diagnoses and treatment of the disorder in girls. Furthermore, symptoms of inattention and disorganization are more likely to be revealed in a more structured educational environment and therefore might lead to a higher age of onset and to less frequent diagnosis in early childhood. These gender differences are mostly given as an explanation for the substantial difference in the male–female ratio between community samples (3:1) and clinical samples (6–9:1).

However, this explanation is not fully satisfactory. The observed gender differences are of rather small size (cf. Gaub & Carlson, 1997; Gershon, 2002). Furthermore, a recent study by Biederman et al. (2005) failed to find any gender differences in the expression of ADHD between boys and girls. Therefore, it seems unlikely that gender differences in the expression of ADHD can fully account for the fact that boys with ADHD receive treatment two to three times more often than girls with ADHD.

However, there is another possible explanation for this phenomenon. It could be that therapists tend to diagnose ADHD in boys more readily than in girls, even if they exhibit the same symptom picture. There is strong evidence that therapists in their clinical routine do not strictly adhere to the diagnostic criteria of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM-IV*) or the International Classification of Diseases (10th rev.; ICD-10; Morey & Ochoa, 1989). Instead, their clinical judgment is affected by certain heuristics and biases. Tversky and Kahneman (1974) showed that people often use certain rules of thumb such as the representativeness heuristic. This means that they base their categorical judgments mostly on principles of similarity (Tversky & Kahneman, 1974) without taking other relevant information, such as the base rate, into account. If a certain object—for instance, an apple—is considered to be similar to the prototype of a certain category, one expects the object to belong to the category in question. Although this is a reasonable approach in daily life, it is not appropriate for diagnosing mental disorders. Clear diagnostic criteria are available for a reliable and comprehensive diagnosis. Nevertheless, several studies have shown that heuristic principles are frequently applied in the context of mental disorder diagnosis. For example, in their well-known study, Langer and Abelson (1974) showed that a videotaped person was rated significantly more disturbed if he was labeled a *patient* rather than a *job applicant*. Subsequent studies showed the influence of several

other factors, such as a patient's gender and race, on the diagnosis (Broverman, Broverman, Clarkson, Rosenkrantz, & Vogel, 1970; Molinari, Ames, & Essa, 1994; Mukherjee, Shukla, Woodlee, Rosen, & Olarte, 1983; Pavkov, Lewis, & Lyons, 1989). Furthermore, several studies found that therapists do not weigh all diagnostic criteria equally, even though this is a requirement in both the *DSM-IV* and the ICD-10. Instead, they weigh diagnostic criteria according to their subjective assumptions about the disorder (Bruchmüller & Meyer, 2009; Kim & Ahn, 2002; Meyer & Meyer, 2009; Schmidt, Salas, Bernert, & Schatschneider, 2005). For example, Kim and Ahn (2002) showed that clinicians were more likely to diagnose a disorder in a hypothetical patient if the patient had causally central rather than causally peripheral symptoms according to the clinicians' personal theory of the disorder. Similarly, in a study of Schmid et al. (2005) in the field of agoraphobia diagnosis, clinicians were asked to rate hypothetical patients with symptom profiles emphasizing each one of three nosological sufficient criteria (avoidance, use of companions, or endurance of situations despite distress). As expected, clinicians weighted the criteria differently. Avoidance was most likely to produce a diagnosis, even if, according to *DSM-IV*, each of the three criteria is equally sufficient for a diagnosis. This shows that clinical diagnosis is also affected by heuristic reasoning.

Furthermore, there is evidence that clinical assessment of child behavior is often biased in the broader sense. For example, several studies showed that the perception of a child also varies as a function of maternal psychopathology (Davé, Nazareth, Sherr, & Senior, 2005). Depressed mothers have been shown to have a more negative perception of their children and are therefore likely to overreport problematic behavior in their children (Chilcoat & Breslau, 1997; Najman et al., 2000). This phenomenon has been named the *depression distortion bias* (Murray & Cooper, 1997).

It is likely that comparable biases and heuristic principles are also applied in the context of ADHD diagnosis. As mentioned earlier, boys are more frequently affected by ADHD than are girls in representative population samples. Therefore, a boy with certain ADHD symptoms could be seen as more representative of the "typical ADHD child" than a girl with the same ADHD symptoms and therefore receive a diagnosis of ADHD more easily. The use of heuristics could also be a possible explanation for a potential overdiagnosis of ADHD. If therapists use a prototypical approach in diagnosis, they might give more weight to the core symptoms of a disorder and might more readily overlook certain exclusion criteria because they are not that prominent. This could, in turn, result in overdiagnosis.

Taken together, although a diagnosis should be based on *DSM-IV* or ICD-10 criteria, the findings described above suggest that diagnosis might be considerably influenced by heuristics in clinical routine. If this was the case, this would have major impact on the quality of treatment, as an accurate diagnosis is an important precondition for disorder-specific efficacious treatment (e.g., Basco et al., 2000; Shear et al., 2000).

In our study, we wanted to assess whether clinicians are influenced by the representativeness heuristic (Tversky & Kahneman, 1974) rather than using the *DSM-IV*/ICD-10 criteria required for the specific diagnosis. On the basis of Tversky and Kahneman's (1974) theory of representativeness heuristic, we hypothesized that therapists base their ADHD diagnoses mainly on the most prominent symptom criteria and therefore diagnose ADHD in cases for

which an ADHD diagnosis would be inappropriate according to *DSM-IV* and ICD-10. This could happen if less prominent criteria are missing or the symptoms are a part of another disorder that has some symptom overlap with ADHD (e.g., oppositional defiant disorder, generalized anxiety disorder [GAD]). This approach would result in overdiagnosis of ADHD. Because boys represent the prototypical gender for ADHD, our second hypothesis was that clinicians diagnose ADHD more readily in boys than in girls, even if the symptoms exhibited are the same. To assess these hypotheses, we sent case vignettes to 1,000 therapists in Germany. The vignettes either fulfilled or did not fulfill ADHD criteria and described either a boy or a girl. We expected that ADHD would be diagnosed in vignettes where necessary diagnostic criteria were not fulfilled. Furthermore, we expected that ADHD would be diagnosed more frequently in the vignettes featuring a boy compared with vignettes featuring a girl.

Method

Participants and Procedure

Participants were 473 psychotherapists specializing in children and adolescents. We selected potential participants by drawing a random sample of 1,000 out of an official register of all therapists specializing in children and adolescents working in the four German states of Baden-Württemberg, Bavaria, Lower Saxony, and Hesse. This register included all psychologists, psychiatrists, and social workers licensed to practice child and youth psychotherapy. Each of the therapists received by post a cover letter, a case vignette, and a questionnaire. We created eight different versions of the case vignette (see below). Therefore, the sample of 1,000 was divided into eight subgroups of 125 therapists each who received the same vignette. The eight groups were balanced for gender and professional occupation (psychologist, psychiatrist, or social worker). If therapists did not return the questionnaire within 6 weeks, a reminder letter was sent out. As an incentive for participation, the therapists could be informed about the results of the study. The sample size of 1,000 was chosen on the basis of a power analysis assuming a response rate of about 40% and a medium effect size. The data collection took place between September and November of 2009.

The effective response rate was 47% with 348 (35%) therapists responding to the initial letter and a further 125 (12%) responding to the reminder letter. We received a total of 473 responses. For organizational reasons, it was not possible to assess whether some letters were returned because of changes in address, deceased respondents, or other reasons. As this might have been the case, it is likely that the response rate of 47% is an underestimation. Essential data were missing in 10 cases; these participants were excluded and we ended up with a final sample of 463 respondents. We compared the age and gender distribution of our sample with the official data of all German child and adolescent psychotherapists published by the German national organization of psychotherapists (*Bundespsychotherapeutenkammer*). Means of our sample (M age = 53.25 years, 68.4% women) were very close to the population data (M = 53.54 years, 69.3% women). This supports the representativeness of our sample for the population of child and adolescent psychotherapists in Germany. The response rate did not differ in respect to the type of vignette (range: 54–65 per

vignette), $\chi^2(7, N = 473) = 2.214, p = .947$. Most analyses are based on the 463 responses, although the sample size might vary because of individual missing answers.

To confirm that randomization was achieved, we compared the groups of therapists who received the same vignette with therapists' characteristics. There were no significant differences between the vignette groups with respect to therapist age, $F(7, 440) = 0.650, p = .714$; theoretical orientation (psychodynamic vs. cognitive behavioral vs. both), $\chi^2(14, N = 463) = 19.603, p = .143$; years of job experience, $F(7, 453) = 0.965, p = .456$; and self-reported *DSM-IV/ICD-10* knowledge, $F(7, 449) = 0.714, p = .660$.

Of the respondents, 68.4% were women. There was no significant difference in the gender distribution between the initial random sample and the participants, $\chi^2(1, N = 455) = 0.43, p = .836$. Of the therapists, 81 (17.8%) were psychiatrists, 248 (54.5%) were psychologists, and 126 (27.7%) were social workers. There was also no significant difference between the initial random sample and the participants in terms of their professional occupation, $\chi^2(2, N = 455) = 0.473, p = .789$. The mean age of the final sample was 53.25 years ($SD = 7.1$). The sample had an average of 18 years of job experience ($SD = 8.2$).

Material Sent to the Participants

Cover letter. In the cover letter, therapists were asked to read the case vignette and to answer the questionnaire. To enhance ecological validity, we included in the cover letter advice noting that despite the difference between a case story and a real setting, therapists were nevertheless being asked to treat this case story as a real case.

Case vignettes. We designed four different case vignettes. The vignettes were constructed on the basis of the *DSM-IV* and ICD-10 criteria of ADHD. Two versions of the ICD-10 exist: the ICD-10 research criteria and the ICD-10 clinical guidelines. The ICD-10 research criteria for ADHD are nearly identical with the *DSM-IV* criteria for ADHD with the difference that the ICD-10 research criteria require symptoms of both inattentiveness and hyperactivity. Therefore, an ADHD diagnosis according to ICD-10 research criteria is identical with a *DSM-IV* diagnosis of ADHD combined type. They both require the following criteria: (a) six symptoms of inattention and six symptoms of hyperactivity/impulsivity, (b) onset before 7 years of age, (c) impairment in two or more settings, (d) clinically significant impairment, and (e) symptoms not better accounted for by another disorder. The ICD-10 clinical guidelines are much less operationalized. They contain clinical descriptions of a disorder, and the ICD-10's diagnostic guidelines are more recommendatory in character. However, concerning the diagnosis of ADHD, the ICD-10 clinical guidelines mention all of the aspects that are included in the diagnostic criteria of *DSM-IV* and ICD-10 research criteria. Our vignettes were constructed so that the underlying disorders could be unambiguously diagnosed according to all three manuals. The vignettes varied as follows:

Vignette 1: ADHD fulfilled. The first vignette described a youth with ADHD. The vignette contained all of the information necessary to clearly diagnose ADHD (combined type) according to *DSM-IV*, ICD-10 research criteria, and ICD-10 guidelines. All

five of the required criteria (see above) were mentioned as being fulfilled.

Vignette 2: No ADHD, two criteria missing. The second vignette was identical with the first vignette, except that Criteria b and c of the *DSM-IV* and ICD-10 research criteria for ADHD were explicitly noted to be nonfulfilled. In the case vignette, the symptoms were present only in one setting (school) and the symptoms only appeared after the age of 9 years. For this reason, even though this vignette contained the core symptoms of ADHD (Criterion a), an ADHD diagnosis could not be given according to *DSM-IV* and ICD-10 criteria, because Criteria b and c were not fulfilled. The ICD-10 clinical guidelines also state that “the symptoms should occur in more than one setting” and “the characteristic behavior problems should be of early onset (before age 6 years)” (World Health Organization, 1992, p. 207). Therefore, an ADHD diagnosis could not be given in this vignette when using any of the three manuals.

Vignette 3: No ADHD, three criteria missing. The third vignette was similar to Vignette 2 with the difference that fewer symptoms of inattentiveness and of hyperactivity/impulsivity were present. Vignette 3 mentioned only three symptoms of inattentiveness and two symptoms of hyperactivity/impulsivity. Therefore, Criterion a was also not fulfilled. As in Vignette 2, Criteria b and c were not fulfilled. Therefore, Vignette 3 contained some of the core symptoms of ADHD, but, in accord with *DSM-IV* and both ICD-10 research criteria and clinical guidelines, an ADHD diagnosis could not be given because Criteria a, b, and c were not fulfilled.

Vignette 4: No ADHD, GAD with symptom overlap. The fourth vignette fulfilled the criteria for GAD, according to *DSM-IV* and ICD-10 research criteria and clinical guidelines. As agitation is also a possible symptom of GAD, a certain overlap exists between the *DSM-IV* and ICD-10 criteria of GAD and ADHD. Vignette 4 described symptoms of restlessness, nervousness, and difficulties in concentration that are similar to ADHD

symptoms. However, no other ADHD criteria were mentioned, and the case vignette fulfilled all criteria for a GAD diagnosis.

Gender variation. To test whether the gender of the youth influenced the therapists' diagnostic decision, we formulated all four vignettes in boy and girl versions. The boy was named Leon and the girl was named Lea. All other information in the vignettes was identical. All eight vignettes (the boy and girl versions of all four vignettes) can be found online in the supplemental materials.

Vignette construction and pretest. The vignettes were constructed by all three authors, two of whom (Silvia Schneider and Jürgen Margraf) are experts in mental illness diagnosis in Germany (e.g., Schneider & Margraf, 2006; Schneider, Unnewehr, & Margraf, 2009). The first version of the vignettes was given to 14 diagnosticians (Pretest 1). They all held degrees in psychology, had completed a standardized training course in structured interviewing, and had successfully passed a reliability check. At the time of the study, they were working as diagnosticians in large research projects and in a behavior therapy-oriented university outpatient clinic. Diagnosticians were all unaware of the purpose and hypotheses of this study. They were asked to make a diagnosis for the vignettes on the basis of each of the diagnostic manuals (*DSM-IV*, ICD-10 research criteria, and ICD-10 guidelines). Furthermore, they were asked to indicate their confidence in the correctness of their diagnosis and to note aspects they were struggling with in diagnosing the disorders depicted in the vignettes. The diagnosticians gave the intended diagnosis in 91.7%, 83.3%, 100%, and 88.1% of the cases for Vignettes 1–4, respectively (see Table 1). Confidence ratings can be seen in Table 1.

After this first pretest, the clarity of the vignettes was discussed in a meeting with five experienced researchers and trained diagnosticians; each member of this group held either a master's or a doctoral degree in clinical psychology. On the basis of the feedback from the pretest and the research meeting, we adjusted the vignettes to make them entirely unambiguous. The revised vignettes were then given to four different diagnosticians with the

Table 1
Results of the Vignette Pretest

Task	Vignette 1	Vignette 2	Vignette 3	Vignette 4
Pretest 1				
Diagnosis based on ICD-10 guidelines: Intended diagnoses	100%	50%	100%	92.9%
Average confidence rating	80%	61.25%	81.7%	62.6%
Diagnosis based on ICD-10 research criteria: Intended diagnoses	100%	100%	100%	85.7%
Average confidence rating	85%	72.5%	90%	68.6%
Diagnosis based on <i>DSM-IV</i> criteria: Intended diagnoses	75%	100%	100%	85.7%
Average confidence rating	85%	75%	85.8%	82.1%
Average intended diagnoses	91.7%	83.3%	100%	88.1%
Pretest 2				
Diagnosis based on ICD-10 guidelines: Intended diagnoses	100%	100%	100%	100%
Average confidence rating	84.6%	77.5%	80%	73.8%
Diagnosis based on ICD-10 research criteria: Intended diagnoses	100%	100%	100%	100%
Average confidence rating	86.3%	82.5%	88.8%	82.5%
Diagnosis based on <i>DSM-IV</i> criteria: Intended diagnoses	100%	100%	100%	100%
Average confidence rating	89.6%	83.4%	89.5%	90%
Average intended diagnoses	100%	100%	100%	100%

Note. Pretest 1 consisted of the first version of the vignettes. Fourteen diagnosticians diagnosed two vignettes each: either Vignette 1, 2, or 3 and Vignette 4. Pretest 2 consisted of the revised version of the vignettes. Four diagnosticians each diagnosed all of the vignettes. ICD-10 = International Classification of Diseases, 10th revision; *DSM-IV* = *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition.

same qualifications as above (Pretest 2). The intended diagnosis was given to 100% of the cases for all vignettes. This means that all of the diagnosticians gave a diagnosis of ADHD in Vignette 1, no disorder in Vignettes 2 and 3, and GAD in Vignette 4 (see Table 1). Confidence ratings can be seen in Table 1. After we explained the purpose of the study to the diagnosticians, they made some additional minor suggestions to further remove ambiguity from the vignettes, which we incorporated.

Questionnaire. All therapists received the same questionnaire regardless of their version of the vignette. The questionnaire contained the following sections.

Diagnosis. Psychotherapists were asked to make a diagnosis by indicating the F code (e.g., F90.0) as well as writing down the label of the disorder (e.g., ADHD). We chose the ICD-10 code because the health care system in Germany requires a diagnosis based on ICD-10, and it is therefore more commonly used than *DSM-IV* among psychologists and psychotherapists. However, diagnostic criteria for ADHD are nearly identical in ICD-10 and *DSM-IV*. Therefore, if the diagnostic criteria for the ICD-10 category *F90 Hyperkinetic Disorders* are fulfilled, it is also possible to make a *DSM-IV* diagnosis of ADHD (combined type).

To analyze the diagnoses, we categorized the therapists' answers into two categories: *ADHD diagnosis* and *no ADHD diagnosis*. The therapists' answers were categorized as ADHD diagnosis if they had indicated the F codes 90.0 or 90.1 or had explicitly indicated a diagnosis of ADHD. If a therapist had written "suspected ADHD," the case was not classified as ADHD diagnosis.

Therapists' answers in the category *no ADHD diagnosis* were further classified into four subcategories. These subcategories were *I do not yet have enough information to give a diagnosis (not enough information)*, *no diagnosis for Lea/Leon (no diagnosis)*, *other diagnosis*, and *suspected ADHD*. We developed the categories on the basis of the therapists' answers. The answers of the therapists were categorized independently by three separate raters. Cohen's κ for the agreement of the three raters was between .944 and .958, indicating a reliable categorization.

Treatment recommendation. In addition, we asked the therapists for their treatment recommendations. They could indicate whether they would recommend any treatment at all (yes vs. no), psychotherapeutic treatment (yes vs. no), and medication treatment (yes vs. no).

Therapist attributes. In addition, we collected the therapists' sociodemographic data such as gender, age, and years of job experience. The therapists were also asked to state their main therapeutic approach, that is, cognitive behavioral, psychodynamic, gestalt, client-centered, systemic, or other. We furthermore asked the therapists how helpful they find *DSM-IV/ICD-10* diagnoses for their clinical routine (on a scale from 0 = *not at all* to 100 = *very much*) and asked them to estimate their familiarity with *DSM-IV/ICD-10* (on a scale from 0 = *vaguely familiar* to 100 = *very familiar*).

Statistical Analysis

To investigate whether ADHD is, in fact, overdiagnosed, we first calculated the percentage of ADHD diagnoses in all vignettes. According to the definition of Scituito and Eisenberg (2007), the number of false positive diagnoses should significantly exceed the

number of false negative diagnoses for a disorder to be overdiagnosed. Therefore, we conducted a chi-square test to compare the proportion of false positive diagnoses (diagnosis of ADHD in the non-ADHD vignettes) with the proportion of false negative diagnoses (diagnosis other than ADHD in the ADHD vignette).

To investigate whether the gender of the case vignette had an influence on the diagnosis, we conducted a multiple logistic regression analysis. The diagnosis (ADHD vs. other diagnosis) was the dependent variable and gender of child in case vignette (boy vs. girl) was the predictor variable. To assess whether any additional factors influenced the diagnostic decision, we included the variables gender of therapist, age of therapist, therapeutic approach, years of job experience, and professional occupation (psychiatrist, psychotherapist, or social worker) and included the two interactions Gender of Child in Case Vignette \times Gender of Therapist and Gender of Child in Case Vignette \times Type of Case Vignette (ADHD vs. non-ADHD) as predictors. Tests for multicollinearity were conducted prior to the analysis.

Results

Overdiagnosis of ADHD

We expected that in non-ADHD vignettes, a substantial number of therapists would diagnose ADHD. This expectation was confirmed. Figure 1 shows the percentage of ADHD diagnoses for the various vignettes. Taken together, in all six non-ADHD vignettes (Vignettes 2–4 with Leon and Lea), 16.7% of the therapists diagnosed ADHD (see also Table 2), 57% of the therapists made another diagnosis, and 10.2% stated that they would give no diagnosis at all. Some therapists did not make a definite diagnostic decision, even though they were asked to do so. Of the therapists, 9.9% stated that they did not have enough information to give a diagnosis. Of the therapists, 5.8% noted "suspected ADHD" instead of making a definite diagnosis.

In contrast, 78.9% of therapists diagnosed ADHD in Vignette 1 (ADHD fulfilled), whereas 7% of the therapists gave a diagnosis other than ADHD (see Table 2). A number of therapists did not make a definite diagnostic decision. That is, 9.6% stated that they did not have enough information to give a diagnosis, and a further 4.4% diagnosed "suspected ADHD."

In addition, we compared the rate of false positive diagnoses (diagnosis of ADHD in the non-ADHD vignettes) with the rate of false negative diagnoses (diagnosis other than ADHD in the ADHD vignette). In Vignette 1 (ADHD), eight out of 114 (7.0%) were diagnoses other than ADHD (false negative diagnoses). In contrast, in Vignettes 2–4 (non-ADHD), 57 of 342 (16.7%) were ADHD diagnoses (false positive diagnoses). This difference was significant, $\chi^2(1, N = 456) = 6.513, p = .011$, odds ratio (*OR*) = 2.65, showing significantly more false positive than false negative diagnoses, which indicates an overdiagnosis of ADHD.

Because a substantial number of therapists did not make a definite diagnostic decision (see above), these answers could not yet be categorized as correct, false positive, or false negative diagnoses. In fact, it is unclear whether therapists in the categories *not yet enough information to give a diagnosis* and *suspected ADHD* would end up providing an ADHD diagnosis. Therefore, in a second analysis, we excluded therapists who had not yet made a diagnostic decision. Looking only at the cases with definite diag-

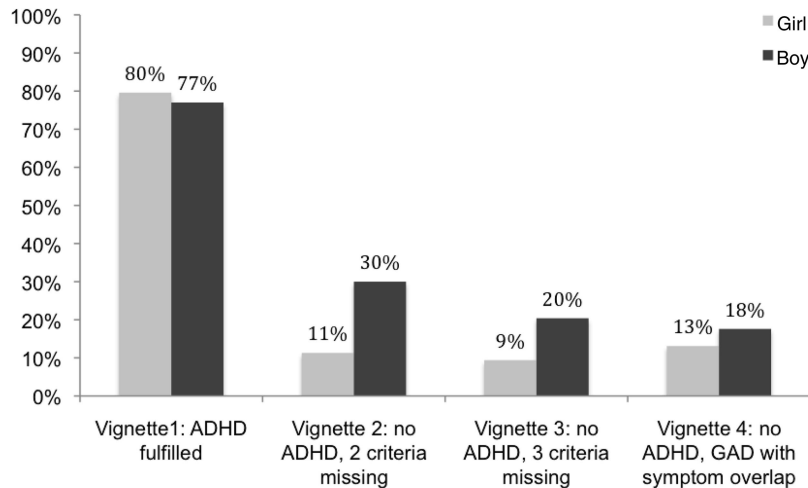


Figure 1. Percentage of attention-deficit/hyperactivity disorder (ADHD) diagnoses for the eight different case vignettes. GAD = generalized anxiety disorder.

nosis, in Vignette 1 (ADHD), eight out of 98 (8.2%) were diagnoses other than ADHD (false negative diagnoses). In contrast, in Vignettes 2–4 (non-ADHD) of the definite diagnoses, 57 of 285 (20%) were ADHD diagnoses (false positive diagnoses). Similar to the results of the first analysis, there were significantly more false positive than false negative diagnoses, $\chi^2(1, N = 386) = 7.060$, $p = .008$, $OR = 2.77$.

All diagnoses other than ADHD were summed up in the category *other diagnosis*. Across all vignettes, the most likely other diagnosis was adjustment disorder. In Vignette 2, out of the 65.3% ($n = 81$) of cases in the category *other diagnoses*, 50.8% ($n = 63$) were diagnoses of adjustment disorder. In Vignette 3, among the 29.0% in the category *other diagnoses*, 14.0% were the diagnosis of adjustment disorder. The rest of the answers in this category

Table 2
Diagnoses Given by Therapists in the Different Case Vignettes

Diagnosis	1		2		3		4		Sum 2–4	
	N	%	N	%	N	%	N	%	N	%
Results for vignettes featuring a girl										
ADHD	43	81.1	7	11.1	5	9.4	8	13.1	20	11.3
Other diagnosis	4	7.5	46	73.0	15	28.3	45	73.7	106	59.9
No diagnosis	0	0	1	1.6	16	30.2	0	0	17	9.6
Not enough information	3	5.7	4	6.3	12	22.6	6	9.8	22	12.4
Suspected ADHD	3	5.7	5	7.9	5	9.4	2	3.3	12	6.8
Sum	53	100	63	100	53	100	61	100	177	100
Results for the vignettes featuring a boy										
ADHD	47	77.0	18	29.5	11	20.4	8	16.0	37	22.4
Other diagnosis	4	6.6	35	57.4	16	29.7	39	78.0	90	54.5
No diagnosis	0	0	1	1.6	17	31.5	0	0	18	10.9
Not enough information	8	13.1	3	4.9	7	13.0	2	4.0	12	7.3
Suspected ADHD	2	3.3	4	6.6	3	5.6	1	2.0	8	4.8
Sum	61	100	61	100	54	100	50	100	165	100
Sum of the results of the vignettes featuring a girl and vignettes featuring a boy										
ADHD	90	78.9	25	20.2	16	15.0	16	14.4	57	16.7
Other diagnosis	8	7.0	81	65.3	31	29.0	84	75.6	196	57.3
No diagnosis	0	0	2	1.6	33	30.8	0	0	35	10.2
Not enough information	11	9.6	7	5.6	19	17.8	8	7.2	34	9.9
Suspected ADHD	5	4.4	9	7.3	8	7.5	3	2.7	20	5.8
Sum	114	100	124	100	107	100	111	100	342	100

Note. ADHD = attention-deficit/hyperactivity disorder; No diagnosis = answers of therapists in the category “no diagnosis for Lea/Leon”; Not enough information = answers of therapists in the category “I have not enough information for already making a diagnosis.”

referred mostly to unspecific diagnoses such as disorder not otherwise specified. In Vignette 4, the most likely diagnosis within the category *other diagnosis* (75.6%) was GAD (38.7%). The rest of the answers in this category were mostly from the anxiety disorder spectrum.

A comparison of the ADHD diagnoses in the three non-ADHD vignettes reveals that Vignette 2 had the highest likelihood of an ADHD diagnosis. On average, 20.2% of therapists diagnosed ADHD even though two necessary diagnostic criteria were not fulfilled. In Vignettes 3 and 4, 15% and 14.4% of clinicians, respectively, still diagnosed ADHD.

Diagnosis of ADHD in the Girl Versus Boy Vignettes

We expected that ADHD would be diagnosed more frequently in the boy vignettes than in the girl vignettes. This hypothesis was confirmed by our data. We performed a logistic regression analysis to determine the influence of vignette gender, as well as other predictors, on the diagnosis of ADHD. Table 3 presents the odds ratio estimates for the logistic regression model as well as their significance. The model chi-square was statistically significant, $\chi^2(12, N = 420) = 120.688, p < .0001$, indicating that the equation with the predictors fits the data significantly better than the equation without the predictors. Controlling for all other predictors, we found that the gender of the child in the vignette was a significant predictor for diagnosis. Compared with the odds for the girl vignette, the odds of clinicians making an ADHD diagnosis in the boy vignette was more than twice as high ($OR = 2.66, p = .034$). Figure 1 shows that when adding up the diagnoses in the three non-ADHD vignettes (i.e., Vignettes 2–4), about twice as many ADHD diagnoses were given in the boy condition compared with the girl condition. However, Figure 1 reveals that this difference occurred only in Vignettes 2–4. In Vignette 1 (ADHD fulfilled), there is no clear difference in diagnoses between the boy and girl vignettes (77% vs. 80%, respectively). We separately

analyzed the proportion of false negative and false positive diagnoses for the boy and girl vignettes. For the girl vignette, there was no significant difference between the proportion of false positive (11.3%) and false negative (7.5%) diagnoses, $\chi^2(1, N = 230) = 0.614, ns$. In the boy vignette, there were significantly more false positive (21.8%) than false negative (6.6%) diagnoses, $\chi^2(1, N = 226) = 7.12, p = .008$.

Influence of Therapists' Characteristics on Diagnostic Decision

We also wanted to assess whether other attributes of the therapists were related to diagnosis. Therefore, we tested the variables gender of the therapist (man vs. woman), professional occupation (psychologist, psychiatrist, or social worker), theoretical orientation (cognitive behavioral vs. psychodynamic), age, and years of job experience in the regression model. We additionally sought to determine whether interaction existed between the gender of the therapist and the gender of the child described in the case vignette. Therefore, we included the interaction Gender of Therapist \times Gender of Child in Case Vignette. As additional predictors, we included therapists' estimation of their *DSM-IV/ICD-10* knowledge and their estimation of the helpfulness of *DSM-IV/ICD-10* for their clinical practice. Table 3 shows the results of the regression analysis. Among the variables tested, only the factor gender of therapist had a significant influence on the given diagnosis. The odds ratio ($OR = 0.267, p = .001$) indicates that across all case vignettes, men were significantly more likely to make an ADHD diagnosis than were women. However, the interaction Gender of Therapist \times Gender of Child in Case Vignette was not statistically significant ($p = .157$).

Looking at the proportion of false positive versus false negative diagnosis for male and female therapists separately, there was no significant difference between the proportion of false positive (12.3%) and false negative (9.6%) diagnoses, $\chi^2(1, N = 309) =$

Table 3
Results for the Regression Model Predicting Attention-Deficit/Hyperactivity Disorder Diagnosis Versus Other Diagnoses

Variable	B	SE	Odds ratio	p
Gender vignette: male vs. female	0.980	0.462	2.66	.034
Gender therapist: female vs. male	-1.322	0.409	0.267	.001
Gender vignette by gender therapist	0.824	0.582	2.279	.157
Age of therapist	-0.018	0.027	0.982	.500
Professional occupation				.180
Psychiatrist vs. psychologist, social worker	0.635	0.451	1.888	.159
Psychologist vs. psychiatrist, social worker	0.657	0.368	1.928	.074
Theoretical orientation				.514
Cognitive behavioral vs. psychodynamic, both	0.859	0.852	2.362	.313
Psychodynamic vs. cognitive behavioral, both	0.586	0.843	1.797	.487
Years of job experience	0.021	0.023	1.021	.375
Helpfulness <i>DSM-IV/ICD-10</i> diagnoses	0.010	0.007	1.010	.128
Familiarity with <i>DSM-IV/ICD-10</i>	-0.006	0.009	0.994	.494
Version of case vignette				.000
1 vs. 2, 3, 4	3.544	0.416	34.601	.000
2 vs. 1, 3, 4	0.446	0.402	1.562	.268
3 vs. 1, 2, 4	0.089	0.433	1.093	.838
Constant	-1.565	1.724	0.209	.364

Note. *DSM-IV* = *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition; *ICD-10* = International Classification of Diseases, 10th revision.

0.454, *ns*, for female therapists. This was also the case in looking at the girl and boy vignettes separately: For girl vignettes, 10.7% false positive versus 11.8% false negative diagnoses, $\chi^2(1, N = 156) = 0.034$, *ns*; for boy vignettes, 14.2% false positive versus 7.5% false negative diagnoses, $\chi^2(1, N = 153) = 1.205$, *ns*.

For male therapists, on the contrary, there were significantly more false positive (25.7%) than false negative (2.6%) diagnoses, $\chi^2(1, N = 143) = 9.441$, $p = .002$, $OR = 12.82$. A closer look at the boy and girl vignettes separately shows that in the girl vignettes, the proportion of false positive (13.0%) versus false negative diagnoses (0%) did not reach statistical significance, $\chi^2(1, N = 72) = 2.585$, *ns*. In contrast, in the boy vignettes, there were significantly more false positive (39.2%) than false negative (5.3%) diagnoses, $\chi^2(1, N = 71) = 8.075$, $p = .004$, $OR = 12.20$. As a substantial number of therapists did not make a definite diagnostic decision (see above), in a second analysis, we excluded these therapists. Looking only at the cases with definite diagnoses, the results were very similar to the results reported above: For female therapists, 14.9% false positive versus 11.3% false negative diagnoses, $\chi^2(1, N = 257) = 0.501$, *ns*; in a separate analysis for girl vignettes, 13.4% false positive versus 13.3% false negative diagnoses, $\chi^2(1, N = 127) = 0.000$, *ns*; in a separate analysis for boy vignettes, 16.3% false positive versus 9.4% false negative diagnoses, $\chi^2(1, N = 130) = 0.934$, *ns*. For male therapists, 29.3% false positive versus 2.9% false negative diagnoses, $\chi^2(1, N = 126) = 10.016$, $p = .002$, $OR = 13.70$; in a separate analysis for girl vignettes, 15.6% false positive versus 0% false negative diagnoses, $\chi^2(1, N = 61) = 2.812$, *ns*; in a separate analysis for boy vignettes, 42.6% false positive versus 5.6% false negative diagnoses, $\chi^2(1, N = 65) = 8.146$, $p = .004$, $OR = 12.66$.

Influence of Diagnosis on Treatment Recommendation

To examine whether overdiagnosis of ADHD has an impact on treatment, we assessed therapists' treatment recommendations for the non-ADHD vignettes (Vignettes 2–4). We calculated chi-square tests to compare treatment recommendations of therapists who diagnosed ADHD with treatment recommendations of therapists who did not diagnose ADHD in Vignettes 2–4. Therapists who diagnosed ADHD significantly more frequently recommended medication (33.3% vs. 3.3%), $\chi^2(2, N = 291) = 48.319$, $p = .001$, and psychotherapeutic treatment (92.6% vs. 83.2%), $\chi^2(1, N = 298) = 3.045$, $p = .041$. Therapists who diagnosed ADHD in Vignettes 2–4 also chose the alternative no treatment at all significantly less frequently than did therapists who did not diagnose ADHD (5.4% vs. 16.9%), $\chi^2(1, N = 311) = 4.823$, $p = .014$. This shows that an ADHD diagnosis in Vignettes 2–4 goes along with a higher likelihood of treatment, especially for medication. These results demonstrate that overdiagnosis has a direct impact on both the individual child and society as a whole. Children are unnecessarily treated and thereby exposed to the associated risks while the health care system faces higher treatment costs.

Discussion

In this survey of 473 child and adolescent psychotherapists, we investigated the diagnosis of ADHD with a set of case vignettes either fulfilling or not fulfilling ADHD criteria and describing

either a boy or a girl. We found that about one sixth of the therapists (16.7%) diagnosed ADHD even though relevant criteria were not given, whereas in vignettes providing all necessary diagnosis criteria for ADHD, only around half as many of therapists (7%) gave a diagnosis other than ADHD. This result supports our first hypothesis of an overdiagnosis of ADHD. The higher rate of false positive compared with false negative diagnoses shows that there is not just a lack of reliability in ADHD diagnosis, as this would result in an equal number of false positive and false negative diagnoses. Our second hypothesis that boys more easily receive an ADHD diagnosis than girls even if they exhibit the same symptoms was also supported. Furthermore, male therapists were more likely to make an ADHD diagnosis than were female therapists. If therapists made an ADHD diagnosis in vignettes where ADHD criteria were not completely fulfilled, they also more frequently recommended medication for treatment.

In Vignette 2, more than 20% of the therapists diagnosed ADHD even though two relevant diagnostic criteria were not fulfilled. It would be interesting to assess the percentage of ADHD diagnoses in a case vignette where only one criterion is missing. Our results suggest that in such a vignette, the proportion of ADHD diagnoses would possibly be even higher than the observed 20%. Our results might therefore even underestimate the proportion of overdiagnosis of ADHD in clinical practice. Furthermore, across all vignettes, 9.9% of the therapists stated that on the basis of the information given in the case vignette, they would not yet make a diagnosis. An additional 5.5% of the therapists diagnosed suspected ADHD. As some of the therapists might possibly end up making a diagnosis of ADHD after a further evaluation, the real percentage of ADHD diagnoses would be even higher than the numbers reported in this study. This once again suggests that our results underestimate the proportion of overdiagnosis of ADHD.

The results concerning Vignette 4 show that a distinct disorder with symptoms overlapping with ADHD is somewhat likely to be classified as ADHD. This agrees with the results of Abikoff, Courtney, Pelham, and Koplewicz (1993) and Jackson and King (2004). Jackson and King showed teachers a video with either a boy or a girl presenting the same oppositional defiant behavior. They then asked teachers, among other questions, to rate the hyperactivity of the presented child. The boy tape received significantly higher hyperactivity ratings than the girl tape did, even though both tapes showed oppositional defiant behavior. Abikoff et al. found similar results in their comparable study. These findings point in the same direction as our results. It is interesting that both in the study of Jackson and King and in our study, this effect occurred mainly in the boy condition. This once again shows the influence of gender on the diagnosis of ADHD.

All in all, our results are consistent with previous studies suggesting possible overdiagnosis of ADHD (Cotuono, 1993; Desgranges et al., 1995; Wolraich et al., 1990). Nevertheless, in their qualitative review of overdiagnosis of ADHD, Sciotto and Eisenberg (2007) concluded that there does not appear to be sufficient justification for the definite conclusion that ADHD is systematically overdiagnosed: "No studies [exist] that compare the diagnoses being given in actual practice to the diagnoses that should have been given based on standardized comprehensive assessments" (p. 107). Our study fills this gap as it systematically assesses whether therapists in clinical practice also diagnose ADHD in cases where necessary diagnostic criteria are not fulfilled. A further strength of

this study is that groups were assigned randomly and were balanced in terms of therapists' gender and professional occupation. Additionally, post hoc analyses were performed to check randomization and control for possible confounding factors, such as therapists' age and other characteristics. Therefore, differences in diagnosis can be attributed to the experimental variation of diagnostic criteria in the vignettes.

As the percentage of false positive ADHD diagnoses was quite high in our study, one might question the adequacy of the diagnostic criteria for ADHD postulated by the *DSM-IV* and *ICD-10*. Concern has been raised in current literature that that the existing *DSM-IV* or *ICD-10* criteria for ADHD might be too conservative or restrictive. However, if clinicians are influenced by heuristics and biases, instead of adhering clearly to diagnostic criteria, this problem cannot be solved by a change in the diagnostic criteria. Therefore, our results suggest that in addition to a potential discussion about the adequacy of diagnostic criteria, the major objective should be the avoidance of the influence of heuristics and biases in psychodiagnosis.

Nevertheless, it is important to keep in mind that diagnostic criteria are constantly being developed. In the upcoming changes for the *DSM-5*, dimensional elements will play a much more important role in diagnosis. This is in line with the findings of Pelham, Fabiano, and Massetti (2005), who showed that the dimension of functional impairment in ADHD is clearly linked to motivation for treatment, even more than core ADHD symptoms. This development suggests that an improvement of diagnostic criteria—in this case, a higher emphasis on impairment—might also lead to a more valid diagnosis and, it is hoped, to a decrease in misdiagnoses.

Concerning the influence of gender on diagnosis, we found that the likelihood for an ADHD diagnosis was clearly higher in the boy vignettes than in the girl vignettes. It is interesting that this gender difference appeared only in the non-ADHD vignettes. Furthermore, significantly more false positive than false negative diagnoses were found only for the boy vignettes. This suggests that boys are more readily overdiagnosed with ADHD than girls are. Our data suggest that if ADHD criteria are all fulfilled, both boys and girls are likely to receive an ADHD diagnosis. In contrast, if only a few ADHD symptoms are present but several diagnostic criteria are missing, boys are much more likely than girls to nevertheless be diagnosed with ADHD. This is an interesting finding, especially as it is frequently assumed in the literature that girls with ADHD are likely to be underidentified with ADHD rather than boys being overidentified with ADHD (e.g., Biederman et al., 2002; Gaub & Carlson, 1997; Sciotto & Eisenberg, 2007). In the literature, it is often argued that girls with ADHD are less likely to have learning disabilities and to show problems at school, in comparison to boys with ADHD (Biederman et al., 2002). Furthermore, girls with ADHD tend to exhibit lower levels of disruptive behavior and higher levels of inattentiveness and internalizing symptoms than do boys with ADHD (Biederman et al., 2005; Gaub & Carlson, 1997; Gershon, 2002). As these symptoms are less likely to disrupt the classroom, it is often assumed that girls with ADHD are likely to be overlooked (e.g., Sciotto & Eisenberg, 2007). In contrast, our data suggest that gender difference in diagnosis of ADHD is caused by an overdiagnosis of ADHD in boys in addition to a potential underdiagnosis in girls. Our results concerning the gender difference provide a reasonable alternative

explanation for the widely discussed differences in the male-to-female ratio in population and clinical samples. Despite a possible difference in ADHD symptoms exhibited by boys and girls, our study clearly shows that exactly the same symptom picture is evaluated differently in boys than in girls.

In our study, male therapists were more likely than female therapists to make an ADHD diagnosis. Furthermore, the effect of patient gender was bigger in male therapists than in their female colleagues. These findings suggest that male therapists are perhaps more likely to be influenced by heuristics or biases in the diagnostic process. However, this was an unexpected finding with no obvious explanation. Researchers should conduct further studies to try to replicate this finding.

Additionally, our study showed that overdiagnosis of ADHD resulted in a more frequent recommendation for medication. This agrees with the findings of Angold, Erkanli, Egger, and Costello (2000), who, in the Great Smoky Mountain Study, found that the majority of children receiving stimulant treatment fell far below the threshold of an ADHD diagnosis. As treatment with stimulants bears the risk of adverse side effects, these findings make clear that misdiagnoses can have a direct and negative impact on patients' treatment. Conversely, about one fourth of the children who received an ADHD diagnosis in the study of Angold et al. (2000) did not receive stimulant treatment. These findings once more stress the importance of a thorough diagnosis as a basis for medication, as both over- and underdiagnosis of ADHD can result in harmful consequences for the patient. Thus, from a public health point of view, our data show that misdiagnoses and the resulting higher treatment costs negatively affect the health care system and society as a whole.

As this study specifically focuses on ADHD, it is not clear how specific the problem of false positive diagnoses is to ADHD in particular. It could be that the same rate of overdiagnosis applies to other disorders as well. In fact, existing studies suggest potential overdiagnoses for other disorders such as schizophrenia or dissociative identity disorder (Lipton & Simon, 1985; Spanos, 1994). However, as these studies were conducted several years ago and more recent studies do not exist, it is unclear whether these disorders still have the tendency to be overdiagnosed. Additionally, there is evidence of the underdiagnosis of bipolar disorder (Bruchmüller & Meyer, 2009). Furthermore, the results of the present study suggest that GAD might be underdiagnosed, as in the GAD vignette, only 38.7% of the therapists diagnosed GAD. All in all, very few studies explicitly deal with the topic of over- or underdiagnosis of specific disorders. However, the studies cited above show that besides ADHD, other disorders are over- or underdiagnosed. Therefore, the number of misdiagnoses in clinical practice might, in fact, be remarkably high. This highlights once more the importance of a thorough and reliable diagnosis. More research is needed to shed light on the biases in specific diagnoses in clinical practice.

As a limitation of our study, one might question the validity of diagnostic and treatment decisions based on written case vignettes compared with real-life settings. For example, when dealing with a case vignette, there is no personal responsibility for decisions. A case vignette also does not allow the therapist to gather further information to confirm or discard his or her decision. However, even in this case, therapists should base their diagnostic decision on the *DSM-IV* or *ICD-10* criteria. Therefore, as we made sure

that our case vignettes either mentioned or explicitly neglected some or all relevant diagnostic criteria of ADHD, the diagnoses based on our case vignettes should be comparable or even less difficult to make than a diagnosis in clinical practice. Additionally, to minimize this difference as much as possible, we advised the therapists to treat the case vignettes as if they were real cases. Nevertheless, ecological validity would have been enhanced if the therapists could have dealt with real clients. However, in an interview situation, each psychotherapist might ask for different information, which would make it nearly impossible to keep the provided information constant.

A second limitation of the study concerns the fact that it is unclear how representative German clinicians are of clinicians in other cultural contexts. Therefore, it is unclear whether these results can be generalized to other countries. However, as several studies from different countries point to problems and biases in diagnosis in general (e.g., Langer & Abelsohn, 1974; Molinari et al., 1994; Pavkov et al., 1989), it is likely that such biases in the diagnosis of ADHD also occur in countries outside of Germany.

In conclusion, the results of this study show that not all therapists follow *DSM-IV* and *ICD-10* requirements to base their diagnosis on a thorough evaluation of the relevant diagnostic criteria. Instead, therapists are influenced by a variety of biases and rules of thumb, such as the representativeness heuristic, which lead to overdiagnosis and different evaluation of symptoms in boys and girls. One way to reduce the influence of diagnostic biases would be to establish more compulsory and thorough methodological and diagnostic training of prospective therapists. Only if therapists recognize how easily diagnostic decisions can be biased can they avoid such pitfalls. In addition, our results indicate how important it is to use structured diagnostic interviews and other standardized tools as accepted instruments in clinical practice. However, recent research has shown that structured interviews are not well accepted and are rarely used in clinical routine (Bruchmüller, Margraf, Suppiger, & Schneider, 2011). To benefit from existing treatments and to ensure an optimal outcome for the patient, it is important to strive for an accurate diagnosis, which is based on accepted diagnostic criteria and not biased by other factors.

References

- Abikoff, H., Courtney, M., Pelham, W. E., & Koplewicz, H. S. (1993). Teachers' ratings of disruptive behaviors: The influence of halo effects. *Journal of Abnormal Child Psychology, 21*, 519–533. doi:10.1007/BF00916317
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Angold, A., Erkanli, A., Egger, H. L., & Costello, E. J. (2000). Stimulant treatment for children: A community perspective. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*, 975–984. doi:10.1097/00004583-200008000-00009
- August, G. J., Realmuto, G. M., MacDonald, A. W., Nugent, S. M., & Crosby, R. (1996). Prevalence of ADHD and comorbid disorders among elementary school children screened for disruptive behavior. *Journal of Abnormal Child Psychology, 24*, 571–595. doi:10.1007/BF01670101
- Barkley, R. (Ed.). (2006). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (3rd ed.). New York, NY: Guilford Press.
- Basco, M. R., Bostic, J. Q., Davies, D., Rush, A. J., Witte, B., Hendricks, W., & Barnett, V. (2000). Methods to improve diagnostic accuracy in a community mental health setting. *The American Journal of Psychiatry, 157*, 1599–1605. doi:10.1176/appi.ajp.157.10.1599
- Berry, C. A., Shaywitz, S. E., & Shaywitz, B. A. (1985). Girls with attention deficit disorder: A silent minority? A report on behavioural and cognitive characteristics. *Pediatrics, 76*, 801–809.
- Biederman, J., Kwon, A., Aleardi, M., Chouinard, V.-A., Marino, T., Cole, H., . . . Faraone, S. V. (2005). Absence of gender effects on attention deficit hyperactivity disorder: Findings in nonreferred subjects. *The American Journal of Psychiatry, 162*, 1083–1089. doi:10.1176/appi.ajp.162.6.1083
- Biederman, J., Mick, E., Faraone, S. V., Braaten, E., Doyle, A., Spencer, T., . . . Johnson, M. A. (2002). Influence of gender on attention deficit hyperactivity disorder in children referred to a psychiatric clinic. *The American Journal of Psychiatry, 159*, 36–42. doi:10.1176/appi.ajp.159.1.36
- Broverman, I. K., Broverman, D. M., Clarkson, F. E., Rosenkrantz, P. S., & Vogel, S. R. (1970). Sex-role stereotypes and clinical judgments of mental health. *Journal of Consulting and Clinical Psychology, 34*, 1–7. doi:10.1037/h0028797
- Bruchmüller, K., Margraf, J., Suppiger, A., & Schneider, S. (2011). Popular or unpopular? Therapists' use of structured interviews and their estimation of patient acceptance. *Behavior Therapy, 42*, 634–643. doi:10.1016/j.beth.2011.02.003
- Bruchmüller, K., & Meyer, T. (2009). Diagnostically irrelevant information can affect the likelihood of a diagnosis of bipolar disorder. *Journal of Affective Disorders, 116*, 148–151. doi:10.1016/j.jad.2008.11.018
- Chilcoat, H. D., & Breslau, N. (1997). Does psychiatric history bias mothers' reports? An application of a new analytic approach. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*, 971–979. doi:10.1097/00004583-199707000-00020
- Cotuono, A. J. (1993). The diagnosis of attention deficit hyperactivity disorder (ADHD) in community mental health centers: Where and when. *Psychology in the Schools, 30*, 338–344. doi:10.1002/1520-6807(199310)30:4<338::AID-PITS2310300408>3.0.CO;2-S
- Davé, S., Nazareth, I., Sherr, L., & Senior, R. (2005). The association of paternal mood and infant temperament: A pilot study. *British Journal of Developmental Psychology, 23*, 609–621. doi:10.1348/026151004X22962
- Desgranges, K., Desgranges, L., & Karsky, K. (1995). Attention deficit disorder: Problems with preconceived diagnosis. *Child & Adolescent Social Work Journal, 12*, 3–17. doi:10.1007/BF01876136
- Esser, G., Schmidt, M. H., & Woerner, W. (1990). Epidemiology and course of psychiatric disorders in school-age children: Results of a longitudinal study. *Journal of Child Psychology and Psychiatry, 31*, 243–263. doi:10.1111/j.1469-7610.1990.tb01565.x
- Fergusson, D. M., Horwood, L. J., & Lynskey, M. T. (1993). Prevalence and comorbidity of *DSM-III-R* diagnoses in a birth cohort of 15 year olds. *Journal of the American Academy of Child & Adolescent Psychiatry, 32*, 1127–1134. doi:10.1097/00004583-199311000-00004
- Gaub, M., & Carlson, C. L. (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*, 1036–1045. doi:10.1097/00004583-199708000-00011
- Gershon, J. (2002). A meta-analytic review of gender differences in ADHD. *Journal of Attention Disorders, 5*, 143–154. doi:10.1177/108705470200500302
- Goodyear, P., & Hynd, G. W. (1992). Attention-deficit disorder with (ADD/H) and without (ADD/WO) hyperactivity: Behavioral and neuropsychological differentiation. *Journal of Clinical Child Psychology, 21*, 273–305. doi:10.1207/s15374424jccp2103_9
- Hartung, C. M., Milich, R., Lynam, D. R., & Martin, C. A. (2002). Understanding the relations among gender, disinhibition, and disruptive behavior in adolescents. *Journal of Abnormal Psychology, 111*, 659–664. doi:10.1037/0021-843X.111.4.659

- Jackson, D. A., & King, A. R. (2004). Gender differences in the effects of oppositional behavior on teacher ratings of ADHD symptoms. *Journal of Abnormal Child Psychology*, *32*, 215–224. doi:10.1023/B:JACP.0000019772.71251.ff
- Kashani, J. H., Orvaschel, H., Rosenberg, T. K., & Reid, J. C. (1989). Psychopathology in a community sample of children and adolescents: A developmental perspective. *Journal of the American Academy of Child & Adolescent Psychiatry*, *28*, 701–706. doi:10.1097/00004583-198909000-00010
- Kim, N. S., & Ahn, W. (2002). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology: General*, *131*, 451–476. doi:10.1037/0096-3445.131.4.451
- Lahey, B. B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., Hynd, G. W., . . . Shaffer, D. (1994). *DSM-IV* field trials for attention deficit hyperactivity disorder in children and adolescents. *The American Journal of Psychiatry*, *151*, 1673–1685.
- Langer, E. J., & Abelson, R. P. (1974). A patient by any other name . . . : Clinical group difference in labeling bias. *Journal of Consulting and Clinical Psychology*, *42*, 4–9. doi:10.1037/h0036054
- Lipton, A. A., & Simon, F. S. (1985). Psychiatric diagnosis in a state hospital: Manhattan State revisited. *Hospital & Community Psychiatry*, *36*, 368–373.
- Meyer, F., & Meyer, T. D. (2009). The misdiagnosis of bipolar disorder as a psychotic disorder: Some of its causes and their influence on therapy. *Journal of Affective Disorders*, *112*, 174–183. doi:10.1016/j.jad.2008.04.022
- Molinari, V., Ames, A., & Essa, M. (1994). Prevalence of personality disorders in two geropsychiatric inpatient units. *Journal of Geriatric Psychiatry and Neurology*, *7*, 209–215.
- Morey, L. C., & Ochoa, E. S. (1989). An investigation of adherence to diagnostic criteria: Clinical diagnosis of the *DSM-III* personality disorders. *Journal of Personality Disorders*, *3*, 180–192. doi:10.1521/pedi.1989.3.3.180
- Mukherjee, S., Shukla, S., Woodle, J., Rosen, A. M., & Olarte, S. (1983). Misdiagnosis of schizophrenia in bipolar patients: A multiethnic comparison. *The American Journal of Psychiatry*, *140*, 1571–1574.
- Murray, L., & Cooper, P. (1997). *Postpartum depression and child development*. New York, NY: Guilford Press.
- Najman, J. M., Williams, G. M., Nikles, J., Spence, S., Bor, W., O'Callaghan, M., . . . Andersen, M. J. (2000). Mothers' mental illness and child behavior problems: Cause–effect association or observation bias? *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 592–602. doi:10.1097/00004583-200005000-00013
- Newcorn, J. H., Halperin, J. M., Jensen, P. S., Abikoff, H. B., Arnold, L. E., Cantwell, D. P., . . . Vitiello, B. (2001). Symptom profiles in children with ADHD: Effects of comorbidity and gender. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*, 137–146. doi:10.1097/00004583-200102000-00008
- Parenting with pills: What you didn't see on the show. (2004, September 28). Available from <http://drphil.com/shows/show/371/>
- Pavkov, T. W., Lewis, D. A., & Lyons, J. S. (1989). Psychiatric diagnoses and racial bias: An empirical investigation. *Professional Psychology: Research and Practice*, *20*, 364–368. doi:10.1037/0735-7028.20.6.364
- Pelham, W. E., Fabiano, G. A., & Massetti, G. M. (2005). Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*, 449–476. doi:10.1207/s15374424jccp3403_5
- Purdie, N., Hattie, J., & Carroll, A. (2002). A review of the research on interventions for attention deficit hyperactivity disorder: What works best? *Review of Educational Research*, *72*, 61–99. doi:10.3102/00346543072001061
- Sandberg, S. (Ed.). (2002). *Hyperactivity and attention disorders of childhood* (2nd ed.). Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511544767
- Schmidt, N. B., Salas, D., Bernert, R., & Schatschneider, C. (2005). Diagnosing agoraphobia in the context of panic disorder: Examining the effect of *DSM-IV* criteria on diagnostic decision-making. *Behaviour Research and Therapy*, *43*, 1219–1229. doi:10.1016/j.brat.2004.09.004
- Schneider, S., & Margraf, J. (2006). *Diagnostisches Interview bei psychischen Störungen (DIPS)* [Diagnostic interview for mental disorders (DIPS)]. Heidelberg, Germany: Springer.
- Schneider, S., Unnewehr, S., & Margraf, J. (2009). *Diagnostisches Interview bei psychischen Störungen im Kindes- und Jugendalter (Kinder-DIPS)* [Diagnostic interview for mental disorders in childhood and adolescence (DIPS for children)]. Heidelberg, Germany: Springer.
- Sciutto, M. J., & Eisenberg, M. (2007). Evaluating the evidence for and against the overdiagnosis of ADHD. *Journal of Attention Disorders*, *11*, 106–113. doi:10.1177/1087054707300094
- Shear, M. K., Greeno, C., Kang, J., Ludewig, D., Frank, E., Swartz, H. A., & Hanekamp, M. (2000). Diagnosis of nonpsychotic patients in community clinics. *The American Journal of Psychiatry*, *157*, 581–587. doi:10.1176/appi.ajp.157.4.581
- Spanos, N. P. (1994). Multiple identity enactments and multiple personality disorder: A sociocognitive perspective. *Psychological Bulletin*, *116*, 143–165. doi:10.1037/0033-2909.116.1.143
- Swanson, J., Arnold, L. E., Kraemer, H., Hechtman, L., Molina, B., Hinshaw S., . . . Wigal, T. (2008a). Evidence, interpretation and qualification from multiple reports of long-term outcomes in the multimodal treatment study of children with ADHD (MTA). Part I: Executive summary. *Journal of Attention Disorders*, *12*, 4–14. doi:10.1177/1087054708319345
- Swanson, J., Arnold, L. E., Kraemer, H., Hechtman, L., Molina, B., Hinshaw S., . . . Wigal, T. (2008b). Evidence, interpretation and qualification from multiple reports of long-term outcomes in the multimodal treatment study of children with ADHD (MTA). Part II: Supporting details. *Journal of Attention Disorders*, *12*, 15–43. doi:10.1177/1087054708319525
- Tversky, A., & Kahneman, D. (1974, September 27). Judgement under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. doi:10.1126/science.185.4157.1124
- Wolraich, M. L., Lindgren, S., Stromquist, A., Milich, R., Davis, C., & Watson, D. (1990). Stimulant medication use by primary care physicians in the treatment of attention deficit hyperactivity disorder. *Pediatrics*, *86*, 95–101.
- World Health Organization. (1992). *International classification of diseases* (10th ed.). Geneva, Switzerland: Author.

Received June 28, 2010

Revision received September 9, 2011

Accepted November 2, 2011 ■