The Role of Measuring Functional Impairment in Assessing Serious Emotional Disturbance in Children

Glorisa Canino, Ph.D.
Professor, Department of Pediatrics
Director, Behavioral Sciences Research Institute
University of Puerto Rico, School of Medicine

Introduction

The 1993 Federal Register definition of severe emotional disturbance (SED), mandated by Public Law 102-321 defines SED as any children from birth to 18 years who "currently or at any time during the past year have had a diagnosable mental, behavioral or emotional disorder of sufficient duration to meet diagnostic criteria specified in the Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised (DSM-III-R)[1], that resulted in impairment, which substantially interferes with or limits the child's role or functioning in family, school, peer relationships or community activities."[2] More recently, the application of this definition to the new Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM 5)[3] and its possible impact on DSM 5 rates of SED was discussed in two expert panels sponsored by the Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality (SAMHSA).[4, 5] After a detailed analysis of all DSM 5 disorders that were included in the SAMHSA definition, the expert panelists concluded that the new psychiatric classification will have little impact on SED estimates and that instruments for the measurement of psychiatric disorders exist for children from 4 years old and above.

Although the criteria for the measurement of psychiatric disorders and the instrumentation for the measurement of these criteria in children are well established, the decision of what constitutes substantial impairment in social, occupational or other areas of functioning is variable with no consensus in the field as to how to define or best measure this construct. In fact, the SAMHSA expert panelists agreed that adequate research, measurement development and associated publications on impairment associated with mental disorders for children are lacking. Relatively recently, a definition of impairment has been provided by the International Classification of Functioning, Disability and Health for Children and Youth (ICF-

CY)[6] (disability as coined by the ICF-CY). However, there is no established instrument for children and youth based on this definition. This paper will discuss the importance of measuring impairment in any mental health survey, the difficulty in defining and operationalizing substantial impairment as required by Public Law 102-123, the existing measures of impairment focusing on the extent to which they can meet the SED criteria by age (0 to 22), issues related to the informant (parent and child), as well as cultural context and language. The review of the literature will focus on population based studies (as compared to clinical samples) and on studies that used psychometrically sound measures of impairment and mental health problems or disorders. Finally, the manuscript will end with recommendations that can be implemented by SAMHSA based on this review.

How to Define Impairment

Impairment in the field of psychiatry and psychology has been often defined as a person's functioning at school/work, in the neighborhood, and with family and friends. In the case of children, functioning is viewed on a continuum, ranging from the highest level of competence to incompetence in adapting to developmental demands posed by the children's context at home (i.e. getting along with parents and siblings), school (i.e. performance in school and getting along with teachers and classmates), and neighborhood (i.e. relating to community members and peers).[7, 8] The lower level, at which normal adaptation is not possible, is considered functional impairment. Most of the instruments of impairment are based on this, or a similar definition, and precede the most recent definition of impairment or disability provided by the ICF-CY.[6] According to the ICF-CY, disabilities (from now on referred to as impairments) are negative functional outcomes resulting from health conditions (physical and emotional), involving significant deviation from or loss of "normal" or "expected" function. Negative functional

outcomes can occur at the individual level as activity limitations or difficulties the child may have with executing activities, and at the societal level as resulting in restrictions in participation or problems the child may have in typical life situations such as school, or with peers, with his/her family and in the community at large.[6] The ICF-CY cautions clinicians when classifying children as impaired or disabled that limitations of functioning or activity performances may in fact reflect delayed development rather than impairment.[9] In addition, the ICF-CY recognizes that the nature of a child's environment differs from that of adults since the child's opportunities to participate in a particular environment are often controlled by parents. The child/caregiver relationship has to be considered as an important part of the child's context and functioning.[9] As the child ages, school becomes an important context since this is where the child spends most of his/her days. Finally, as adolescence approaches both peers and the community at large become much more salient. All of these environments (school, family, societal) are viewed as a barrier or facilitator of a child's functioning. An important contribution of the ICF-CY is that it provides a framework for the development of assessment measures for children.[10] Thus, the conceptualization of impairment contained in the ICF-CY takes into account the developmental stage of the child, the context in which the child lives (family, school, community) and cautions against confusing developmental delays with impairment.

The main difference between the ICF-CY definition of impairment and the previous definitions is the inclusion of physical conditions and the fact that the impairment is considered an outcome of a health condition and therefore must be independent of the health condition itself. This definition is therefore consonant with the SAMHSA definition that separates impairment from disorder by requiring both a DSM disorder and significant impairment in functioning. However, the requirement of separating impairment from the health condition presents a

limitation when the ICF-CY is used to classify children as impaired who also meet criteria for a DSM IV psychiatric disorder. The DSM IV as well as the most recent revision, DSM 5, requires distress or impairment to establish the threshold for a disorder in order to avoid the over inclusive identification of psychiatric cases. In the absence of clear biological markers or clinically useful measurements of severity, it is not possible to separate normal from pathological symptom expressions. The presence of impairment and distress are used by the DSM IV and 5 to further delineate the normal from pathological symptoms. After the introduction of impairment into the DSM criteria the rates of psychiatric disorders were reduced two to threefold in psychiatric epidemiology studies of children.[11] Thus, the identification of the health condition in DSM IV and 5 is dependent upon the presence of functional impairment and is not separate as required by the ICF-CY definition and the Federal Register definition of SED. In fact, the health condition is determined in part based on the impairment it causes. In spite of this, it is well known that several diagnoses are unrelated to impairment in functioning and meeting criteria for a diagnosis is not a good predictor of impairment.[12, 13]

Why is Important to Measure Impairment in Functioning Separate from Diagnosis?

The Measurement of Impairment is Required by Law

Functional impairment can result from a psychiatric or physical illness, a shared etiology with the illness, or from deficits in functioning that may lead to psychiatric sequelae.[14] Thus, both the measurement of psychiatric disorder and impairment are necessary. Furthermore, Public Law 102-321 mandates that the provision of public funds for mental health services for children can only be provided to the states for children who have severe emotional disturbance (SED), that is, that meet criteria for a DSM disorder and for substantial impairment. In addition, managed care plans purportedly follow criteria established by Public Law 102-321 in order to restrict services to those who in their assessment "really" need them, and thereby contain costs and increase their own profit margins. Thus, the classification of functional impairment associated with psychiatric disorder has become a necessary requirement for reimbursement of mental health services and a necessary criterion for the allocation of resources in the public mental health sector.

Impairment is Important for Predicting Outcome and Service Use

In addition to these important considerations, impairment operates differently from psychiatric disorders in predicting course and outcome. Longitudinal studies of children from the Great Smoky Mountains Study have shown that irrespective of treatment, children with both a diagnosis and impairment are almost four times as likely, over a four year period to be arrested, drop out or be expelled from school as those with only a diagnosis, and more than twice as likely as those who were only functionally impaired.[12] Furthermore, decline in functioning, is one of the most common reasons that mental health services are first sought for children and adolescents.[15-17] Impairment in functioning is more likely to lead parents to take their children

into treatment than a diagnosable psychiatric disorder,[18, 19] and perception of impairment appears to be more significant than diagnosis in predicting service utilization in most epidemiologic studies.[12, 18, 19] Moreover, prospective studies show that it is the level of impairment and not symptoms that predicts adult outcomes.[20] Improvement in functioning is the main diagnostic outcome used for determining the effectiveness of any mental health intervention. In addition, identifying the limitations in functioning or impairment in a child is important information potentially useful for planning and implementing service interventions.[21] Lastly, by applying the impairment criteria by specific disorder, it is possible to determine the degree of impairment by disorder, information that is important for prevention and treatment interventions.

The Measurement of Impairment is Cost-Effective

Because of cost restrictions, detailed diagnostic interviews covering a wide range of psychiatric disorders cannot always be conducted in large epidemiologic surveys. Screening instruments, such as the Strength and Difficulties Questionnaire (SDQ)[22] that have been found to predict psychiatric disorders relatively well, are then used to lower costs. However, most screening instruments have limitations in covering the wide range of disorders that children can have and many cases are likely to remain undetected. For example, the SDQ has been shown to not cover well anxiety and eating disorders.[22] The measurement of impairment can then be a good way to detect children in need of services that otherwise would have gone undetected.[23] This is particularly true given that most impairment measures are relatively short and easy to administer.

Impairment Rates Vary by Instrument, Informant, Sex, Age, Race/Ethnicity, and Culture

Symptom Disagreement by Informant, Sex and Age of Child

There is an ample literature that shows that for the most part different informants (i.e. parents, children, teachers) report differentially on different aspects of information, such as psychiatric symptoms, correlates and risk factors.[24-29] Similar lack of agreement between informants is observed in the report of impairment in functioning, although the literature is more limited (see below). In general, parents tend to be better informants (agree more with clinicians) than children (particularly young ones) of psychiatric symptomatology independent of the assessment measure used or the specific diagnosis ascertained.[30-32] Diagnostic information from the parent is also more predictive of outcomes than information obtained from the youth.[31] However, children 12 years of age and older are considered relatively competent informants of their behavior and because of the need for cost containment, several large epidemiologic surveys have only used child informants when interviewing children 12 and above, and parent informants for children below 12 years of age.[33-35] In this way costs are contained and rates of disorder do not vary by informant given that only one informant is used .

For certain psychiatric disorders such as conduct disorder, anxiety, and affective disorders, children report more symptoms than parents.[32, 36, 37] However, adolescent reports of their general and mental health have been found to be lower than those of their parents.[38] In general, better agreement (although still discrepant) between informants is obtained for observable behaviors (such as externalizing symptoms), than for unobservable symptoms, such as suicidal and depressive symptoms.[32, 39-42] Better agreement between informants is also obtained when a youth's behavior involves the action of a third party, such as courts or teachers,[43] and in more severe cases and clinical samples as compared to community samples.

Children in treatment or with greater severity usually have multiple symptoms and diagnoses that are more persistent and incapacitating[38, 44-46] facilitating more agreement among informants. Informant disagreement has been found to be more likely for males, older children, children who are less symptomatic, and those receiving Medicaid.[47] Externalizing symptom informant disagreement was found to be more likely for female children, older, less impaired and less symptomatic.[47] Longitudinal studies have confirmed that disagreement between parent and child informants grows as the child gets older, particularly for internalizing symptoms.[48] This seems natural given that as the child grows older he/she spends less time with the parents and more at school/work and with friends. Other studies have shown that in different contexts (country) and racial/groups, discrepancy among informants in the report of psychiatric symptoms and behavioral problems exist but the magnitude of the discrepancy varies.[35, 49]

Impairment Disagreement by Informant, Sex, Age and Race/Ethnicity

Disagreement between informants is also reported in other studies according to the severity of impairment, the gender and age of the child and the race/ethnicity of the participant.[13,23, 43, 50-52] The most complete study of why parents and children differ in their reports of impairment was examined by Kramer et al.[50] In this study parents and adolescents 11 to 18 years receiving psychiatric treatment were interviewed with a mental health questionnaire and the Columbia Impairment Scale. Results showed that parents and adolescents were more likely to agree on behaviors that had serious consequences or involved a third party such as arrest, incarceration and school expulsion. Similar results were reported previously.[43] Parents as compared to adolescents were more likely to report impairment in role performance, relationships and prosocial activities but less impairment in school or with legal consequences.[50] The greatest disagreement between parent and adolescent was in the areas of relationships and in

9

cases with less severe impairment. Reasons given by parents and adolescents for their

discrepancies were categorized into three main topics: 1) the parent was unaware of the child's

behavior, 2) parent and adolescent interpreted the behavior differently and 3) they both had

different thresholds for defining a behavior as problematic. Comparisons between parent and

clinician, and parent and adolescent, showed that when there was disagreement between

parent/adolescent, clinicians agreed more frequently with parents. The results on informant

impairment disagreement thus parallel those found for psychiatric symptoms.

An epidemiologic study carried out in Germany found differences in prevalence rates of

impairment by informants, sex and age of the child.[23] Parents as compared to boys reported

more definite and severe impairment as measured by the impact supplement (consists of

impairment and distress items) of the Strength and Difficulties Questionnaire (SDQ).[22] Parents

reported more impairment in boys for longer duration and for school and family related

impairments. For girls, parents reported less definite and severe impairment than the girls

themselves. Girls, nevertheless, reported more mild impact than parents. Overall rank

correlations between parent-reported and child reported impact scores (mild, moderate, severe)

was poor (p=.29) irrespective of sex or age group.

In a study carried out in the United States that also used the SDQ impact supplement,

differences in rates were reported by sex, age, poverty level, health insurance, ethnicity and

marital status.[52] Parents reported almost double rates of impairment/distress in boys than girls as

well as those with Medicaid or other public health insurance, more impairment/distress in older

boys and girls, less in Hispanics as compared to other ethnic/racial groups, and more

impairment/distress in single parent and poor families. In another national U.S. study, parents

reported more impairment in older children regardless of sex, and more impairment in boys.[52]

As with the other studies discussed, the Smokey Mountain study[13] reported that impairment was significantly more prevalent in boys as compared to girls, and boys had more school disabilities while girls had more family disabilities. Impairment was more prevalent among ethnic/racial minorities (Hispanics were not included) especially those with disruptive behavior disorders. The higher rate of impairment by minorities could not be explained by higher rates of psychiatric symptoms. Gender differences were also observed regarding the rates by age and type of impairment (family, educational, peers).

How to Best Combine Informants

Because of the discrepancy in the report of information between informants, prevalence rates of disorders and symptom levels vary depending on who is the informant, parent, teacher or child.[15, 19, 53] But because children behave differently in different contexts (family, school, neighborhood, and friends) and because each informant may provide unique and meaningful information, they are both considered essential informants for the diagnostic process, particularly in assessing younger children. There are two general approaches on how to use the information obtained from more than one informant on the same child. The first approach used by some[15, 54] is to present the rates of disorder or impairment separately. The other approach is to combine data from the informants in order to obtain one prevalence rate. This approach is preferred by policy makers who need to plan for services. Combining data from several informants into one rate can be done by adding the positive diagnoses of all informants (AND rule), or what is considered the OR rule, by which a symptom of a diagnosis is counted as positive if, for example, it is endorsed by either parent or child, so that a diagnosis could be positive even if full criteria were not met by either informant. Testing of these two approaches by Bird et al.[55] for psychiatric disorders and the testing for impairment by Kramer et al.[50] showed the OR rule is a

more valuable information base. This OR rule, results in a higher prevalence of the disorders.[56] The problem arises when one informant, such as the parent, is used for the entire age group (i.e. 0 to 22 years) and only one informant for another age group (i.e. 12 to 22 years). The older age group (12 to 22) will have higher prevalence rates, not only because rates increase with age, but also because in this age group, two informants are used whereby in the younger age group only one informant provided the child information.

In summary, there seemed to be agreement across studies regarding the domains, sex and age variables in which parent and child have relative agreement vs. disagreement given that discrepancies in report of impairment and symptoms are the norm. In most studies disagreement seemed to be greater between parent/child among boys, older children, and in unobservable vs. observable behavior. Agreement was more likely in contexts which involved a third party, clinical samples, and more severe impairments. Across most studies impairment seemed to be more common among boys, particularly at school, minority groups with the exception of Hispanics, children on Medicaid or other public insurance, and single parent families.

The Impact of Culture and Context on the Measurement of Impairment

As stated in the ICF-CY[6], disabilities are negative functional outcomes resulting from health conditions involving significant deviation from or loss of "normal" or "expected" function. For the most part impairment measures are based on behaviors or roles that are normative to a given cultural group or context. Given this contextual definition, one would therefore expect that what is considered normal or expected would vary across different cultural groups. In fact, as reported elsewhere[8] comparison between island Puerto Rican children and U.S. children in the scores of the Social Competence Scale of the Child Behavior Checklist (CBCL)[54] showed lower scores in this instrument among island Puerto Rican children because of their lower scores in role performances such as sports, hobbies, organization or jobs that were not available or part of the culture or context of Puerto Rican children. On the other hand, as part of the Epidemiology of Child and Adolescent Mental Disorders (MECA) study[57, 58] children from San Juan, Puerto Rico as compared to children in three other communities of the U.S., scored as less impaired on the Children Global Assessment Scale (C-GAS) and the Columbia Impairment Scale (CIS), in spite of the fact that prevalence rates of disorders were similar in San Juan as compared to the U.S. sites.[58] Similar results to those obtained in Puerto Rico were reported by Costello et al.[59] for American Indians.

Caution should be exerted when choosing an impairment measure that has been validated in only one cultural context. Measures should be carefully evaluated in the cultural and social context where they are intended to be used to avoid bias and misinterpretation of research results from studies using standard methods developed in a single sociocultural context. The C-GAS, for example, and possibly other measures requiring the interviewer to make a judgement about the extent of functional impairment, may be influenced by the rater's experience, prior knowledge of

the child and expectations of what constitutes impairment and social adaptation, particularly if the interviewer belongs to a different culture from that of the respondent.

Prevalence Rates of Impairment Vary Across Studies

Because there is disagreement between informants, by sex, age, ethnicity, and culture, our review of this topic shows great variability in prevalence rates of impairment. Conceptual, contextual and methodological differences add to the discrepant rates observed. In what follows we review these studies with an emphasis on how the rates of impairment across studies vary and the possible factors related to the differences across studies.

An epidemiologic study carried out in Germany found 17.4% of the children between the ages of 7 to 17 had a borderline, definite, or severe impact score of the SDQ.[23] In this representative sample, 12% of the children or parents reported only impairment/distress with no indication of mental health problems. Six percent of these had reported severe impairment/distress but no mental health problems, suggesting that impairment measures are important in capturing false negatives that may be more common in short screening instruments.

In the United States, the SDQ was also used in the National Health Interview Survey, but only parent informants answered the questionnaire. Reports from the 2003 survey showed 4.8% of children 4 to 17 years with definite to severe impact scores (impairment/distress) as indicated by parental report.[52] Mild or borderline impact scores were not reported in this study possibly accounting for the lower rates in this study as compared to the German study. Differences in the context and culture (U.S. versus Germany) and sample composition of these two studies could have also influenced the variability in rates observed. However, both studies shared a similar instrument and conceptualization of impairment/distress.

The latest Medical Expenditure Panel Survey[51] also used the SDQ for the measurement of psychiatric symptoms but for the measurement of impairment this study used the parent version of the Columbia Impairment Scale (CIS) to evaluate impairment in functioning in the U.S. child population 6 to 17 years of age. The CIS is a global impairment measure that is assessed separate from psychiatric symptoms, although it contains a couple of items that are considered psychiatric symptoms.[11, 60] Reports from 2010 to 2012 showed 10.7% of the child population had severe mental health impairment (score of 16 in the CIS). This rate represented a decrease from 12.8% reported previously in 1996-to 1998. Rates of impairment were lowest among Hispanics. About a third of the children with severe impairment used mental health services and psychotropic medication.

The Smokey Mountain Study of North Carolina, measured impairment (called disability by the authors) among a representative sample of children 9 to 16 years of age. Parent and child informants reported on impairment and mental health although no information was provided as to how data from these two informants were combined to provide impairment rates.[13] The study authors reported a much higher prevalence rate of impairment as compared to the previous three studies,[23, 51, 52] 27.3% of children 9 to 16 years met criteria for impairment using a measure of the Child and Adolescent Psychiatric Assessment (CAPA).[13] Differences in rates may be due in part to the use of a different measure of impairment from the previous studies, the combination of two informants for determining the rates, the fact that severity was not considered in the rate, and the differences in contexts (Smokey Mountains versus national sample). Impairment was defined as a decrement in functioning at school, home and elsewhere that occurs as a consequence of psychiatric symptomatology. This conceptualization is measured with 17 items that are asked only after the parent or child endorses symptoms of a disorder. Thus, the impairment measure

refers only to the impact psychiatric symptoms may have on decrements in functioning. Depression and conduct disorder were the two diagnoses with the highest frequency of impairment.

In conclusion, prevalence rates of impairment, as well as of psychiatric disorders and symptoms, vary across different studies that use different informants, samples, contexts, age and sex group, different measures and different conceptualizations of impairment. But perhaps the most difficult task related to the measurement of impairment is determining both conceptually and empirically what is meant by substantial impairment.

How Impaired must a Child be to be Considered Substantially Impaired?

The decision of what constitutes substantial impairment in social, occupation or other areas of functioning is arbitrary. Severity or substantial impairment is determined depending on the clinician's judgment, the cut off scores of the different measurement instruments, and the population ascertained. Different instruments use different methods for ascertaining severity scores: 1) by comparing the mean scores of impairment between clinical and community samples, 2) by using predictors of severity of a disorder as external criteria against which to determine the sensitivity and specificity of the instrument, and 3) by determining the percent of the sample that scores within a specific range of scores (arbitrary score of 5% or 10% depending on the cost of providing the mental health services and the existent budget). Determining what is considered substantial or severe impairment is often made based on cost effectiveness considerations, research purpose or public health policy.

The review of prevalence rates of impairment discussed above showed how prevalence rates varied depending on the conceptual definition of impairment used by the measure, the cut offs used and the context or sample used. There is a need for further research that can examine

the optimal way in which a severity score can be best determined. There is also a need for a more concrete way of interpreting what is considered in the Federal Register definition of SED "substantial impairment in functioning". An operationalization will aid in the development of cut offs of severity that can better respond to the definition of SED. Given this and other limitations discussed previously, we next review the measures of impairment available at present.

Which Impairment Measure is Best Suited to Comply with the Definition of SED?

Introduction

An important goal of this review is to identify an impairment measure that can or has the potential to meet the Federal Register criteria for the measurement of SED. This review will describe mainly the extent to which the various existent instruments meet the SED criteria and the decisions made by the two review panels sponsored by SAMHSA.

Measures of impairment can be classified into three different types: 1) global impairment measures, 2) domain-specific or multidimensional impairment measures and 3) symptom specific or diagnosis specific impairment measures.[8] There is argument about whether it makes more sense to assess impairment globally or separate from the psychiatric disorder or symptomatology measure used or to tie the impairment to specific symptom syndromes as done by two commonly used psychiatric instruments: the Diagnostic Interview Schedule (DISC)[58] and the Child and Adolescent Psychiatric Assessment (CAPA).[61] The SAMHSA expert panel that met on September 8, 2014[4] recommended the inclusion of a separate global measure of impairment that would be developmentally and culturally appropriate, and that would include items to assess impairment in different contexts such as school, family and community. The expert panel also recommended the exclusion of measures of impairment tied to specific symptom syndromes. Using measures of impairment independent of syndromes has the

17

advantage of avoiding confounding impairment with symptom severity. In addition, as discussed by the panel members, and reported by other investigators,[63] many young children may not reach all the diagnostic criteria required by the DSM but may have substantial impairment in functioning, which puts them at high risk and in need of treatment. Furthermore, there is evidence that parents and children have difficulty making attributions about whether the impairment is caused by a particular symptom syndrome, particularly when the child has high rates of disorder, comorbidity, and presents positive scores on several syndromes.[63] Because of this, we will exclude from our review impairment measures tied to specific symptom syndromes within diagnostic schedules. We are also excluding the Vineland Adaptive Behavior Scales,[64] the Scales of Independent Behavior[65] and the Adaptive Behavior Assessment System[66] since these scales are mainly used for determining independence in living and are therefore more useful for children with mental retardation and pervasive developmental disorders. Services for these children are covered by agencies other than SAMHSA. Impairment measures such as the Child and Adolescent Level of Care Utilization System[67] and the Social Adjustment Inventory for Children and Adolescents (SAICA)[68] are also excluded because they are designed to be administered only by clinicians and are therefore not practical in large surveys.

Our review of global and multi-dimensional scales of impairment will focus on whether the instrument meets the following criteria necessary for fulfilling the Federal Register notice definition for SED and is also suitable for administration in large national surveys:

1. Is developmentally appropriate and can be administered to a wide age range (0 to 22), preferably with different versions available depending on age

2. Is a multidimensional scale or has items that assess specific areas of functioning in school, community, and with peers and family

18

3. Has parent and child versions

4. Has good psychometric properties for U.S. population and preferably have Spanish version with psychometrics

5. Is respondent based, that is, the score is dependent on what the parent and child informs and therefore does not require that an interviewer have prior knowledge of the child

6. Ideally is relatively short so that it can be added to existent national surveys that are not focused on mental health

7. It has cut off scores that determine severity of the impairment

8. It is separate from psychiatric symptomatology

There are two relatively recent comprehensive reviews of impairment measures in the literature.[8,14] Since the last two published reviews of impairment measures, there are few new measures of impairment except for the Functional Impairment Scale for Children and Adolescents (FISCA)[69] and the Child World Health Organization Disability Assessment Scale (C-WHO-DAS). This last measure has no published psychometric properties for the U.S. population but we are reviewing this measure mostly because it is the only one based on the International Classification of Disability for Youth (ICF-CY).[6] As stated above, the ICF-CY offers probably the best definition of the construct of impairment that fully captures the Federal Register criteria for SED. In what follows we discuss in more detail each of the measures of impairment that meet at least some of the criteria necessary for the measurement of SED. Review of the scales is divided in global and multidimensional instruments.

Global Scales of Impairment

Children's Global Assessment Scale (C-GAS): The C-GAS is a unidimensional clinician rated[70] and lay interviewer administered impairment scale.[11] The scale measures the child's functioning in the home, community, school and with friends but does not provide scoring by domain, thus making it a global scale. Cut points for defining severe (<61) to moderate impairment (<71) were developed by comparing children with and without diagnoses. The score of impairment given is interviewer based and therefore dependent on the interviewer's knowledge of the child, so that if administered in a large survey by a lay interviewer it has to be administered after other diagnostic scales, thus reducing the length of administration (5 minutes) to give sufficient information to the interviewer to make the impairment judgement. The instrument has no child age limit and potentially can be administered from 0 to 22 years of age. The interrater and test-retest reliability as well as the concurrent and construct validity of the clinician and lay interviewer child and parent versions have been reported for its English and Spanish version.[11,60,70] Moderate agreement between clinician and lay interviewers supports the use of the lay interviewer C-GAS.[11] The instrument was used and tested as part of the Methods for the MECA study, a large epidemiology study that included communities from the U.S. and Puerto Rico.[57]

Some of the advantages of the instrument are the good psychometric properties for clinician and lay interviewers, and its brevity when administered with a comprehensive psychiatric interview. In addition, the instrument can be administered to a wide age range, has severity cut points and has child and parent versions. The disadvantages of the C-GAS are that it requires knowledge of the child, requires rigorous training of lay interviewers, and some of the criteria for assessing impairment are also psychiatric symptoms.

Columbia Impairment Scale (CIS): The CIS is a brief 13 item global respondent based impairment scale that can be administered by lay interviewers without prior knowledge of the child to assess impairment. The parent and youth versions have been tested for internal consistency as well as construct validity in combined clinic and community samples of the U.S. and Puerto Rico as part of the MECA study.[11, 60] Although the scale contains items of four domains (school, peer, home, social) factor analyses revealed only one overall global domain of impairment, thus, as with the C-GAS, only one global score is provided.[60] Severity cut points were tested and a score of 16 or more (considered severe impairment) identified the greatest numbers of cases, use of services, perceived need of services and other severity indicators. The scale can be administered to parents about children from age 4 to 6, and to children themselves from age 7 to 17. Two of the items refer to psychiatric symptoms. In addition to the MECA study, the CIS has been recently used as part of the Medical Expenditures Panel Study (MEPS) previously described.[51] The advantages of the instrument are that it is brief, has good psychometric properties in English and Spanish versions, can be reliably and validly administered by lay interviewers, has a parent and child version, is administered separately from the mental health measure, has items that cover the different child domains or contexts, and it has severity cut points. The main disadvantage is that it is not designed for children under the age of 4, and that two of the items refer to psychiatric symptoms. Only one cut off score considered severe is offered. As the C-GAS, although it has items that tap several domains of the child's functioning, it functions as a unidimensional scale.

Strengths and Difficulties Questionnaire (SDQ) Impact Score: The SDQ is a short (25 items) screening scale that measures emotional and behavioral problems in children.[22] The instrument can be administered to parents and teachers of children 4 to 16 and to youth 11 to 16

years of age. In addition, the SDQ has an impact supplement that contains 5 items that assess distress and impairment (impact score) chronicity and parental burden. The first out of 5 items in this impact supplement asks the respondent (parent, child or teacher) the following: *"Overall do you think that (child's name) has difficulty in any of the following areas: emotions, concentration, behavior, or being able to get along with others.''* A card handed to respondents lists four response options: *no; yes, minor difficulties; yes, definite difficulties; yes, severe difficulties.* If the respondent answers *yes*, another item asks "*How long have these difficulties been present?*'' If the child had these difficulties for one month or more, the respondent is asked whether the difficulties upset the child (distress item) and the extent to which the difficulties interfered with the child's everyday life in home life, friendships, classroom learning, or leisure activities (impairment). The last question in the impact supplement asks the respondent if the difficulties put a burden on the child, parent, or the family as a whole. Thus, out of these five items in the impact supplement of the SDQ, two are directly related to impairment and one is related to the chronicity of the first impairment item. The impact score consists of these three items plus the distress item. The burden item is scored separately and is not considered part of the impact score.

The 25 item SDQ that measures emotional and behavioral symptoms has been found to be a psychometrically sound instrument for the measurement of overall child mental health problems mostly in European countries.[71-76] Fewer studies have examined the psychometric properties of the impact score. The first psychometric properties of the impact items was reported by Goodman[72] in which the cut off scores of the impact measure were determined by comparing clinical and community samples of British children. The most predictive scores of clinical status were scores of 1 and above: 0 was considered no or minor difficulty, 1 definite

difficulty, and 2 severe difficulty. When comparing the scores on the symptom score of the SDQ with that of the impact score, the latter was significantly better in predicting clinical status although the best predictor was a combination of the symptom and impact scores. The test-retest reliability of the impact score was .63.[72] The internal consistency of the SDQ scales and the impact items as well as the development of normative scoring bands for the U.S. population was analyzed as part of the National Health Interview Survey (NHIS), a 3-year follow up study representative of the U.S. population.[77] Only internal consistency was reported separately for the impact score, and it was found to be excellent (.80) as in the British study (.85).[78] In other analyses of the same NHIS study, the results showed that children who were scored as definite or severe cases in the impact items based on the parent reports were ten times more likely to use or receive mental health services.[52] Similar results were reported by Goodman[78] with British children. In addition, a comparison of children with conduct disorder (CD) alone and CD with callous unemotional symptoms (CD+), a group with greater CD severity, showed significantly higher impact scores in the CD+ group as compared to the CD only group.[79]

Advantages of the SDQ impact items are mostly related to the shortness of the scale, its ability to distinguish between different levels of severity, availability for use by different informants, and the fact that those cases that score in the definite and severe range are more likely to use mental health services, and to meet criteria for a DSM IV disorder. The items also discriminate well between clinical and non-clinical samples. On the other hand, several disadvantages were identified. The impact score or impairment items of the SDQ have not been administered separately from the SDQ. In addition, the SDQ was not designed for children below the age of 4 years and the impact items are not separate from psychiatric symptoms (i.e. concentration problems is a symptom of depression). The impact score also has one distress item.

Multidimensional Impairment Scales

The multidimensional impairment scales are measures that rate a child's functioning in different domains and provide separate scores for each area of functioning. As stated before, children behave differently in different contexts or situations, and it is possible that a child may be impaired in one area and not another. These scales are useful for planning treatment and prevention intervention since they provide information that can be used in the development of an action plan based on the different contexts applicable to the child. As with the global scales, all the dimensional scales discussed mingle psychiatric symptoms with functioning items, preventing full differentiation between these two important constructs. However, scales differ in the amount of symptoms they contain, with the CAFAS as the scale that contains the largest number of psychiatric symptoms and the Columbia Impairment Scale the least.

Child and Adolescent Functional Assessment Scale (CAFAS): The CAFAS rates a child's functioning in eight areas. To rate the caregiving environment, there is a Family/Social Support subscale that rates the youth's functioning based on the limitations encountered in the family's psychosocial resources relative to the youth's needs.[80] The most recent CAFAS version[81, 82] contains 315 items and generates separate scores for each of the role performance areas (school/work, community and home) as well as separate scores for emotions, behaviors towards others, substance use, thinking and moods/self-harm. Each subscale as well as the summary of all scales is scored denoting one of four levels of functioning: minimal or no impairment (0), mild (10), moderate (20) and severe (30). As the C-GAS, the CAFAS impairment scale cannot be administered alone since knowledge of the child is needed. The CAFAS rating is based either on the information obtained on a separate psychiatric interview or on the longer CAFAS interview designed to collect the necessary information to rate the

impairment in functioning. Because of this, although the impairment scale takes around 10 minutes to administer, when the CAFAS interview is added, the time of administration can reach 40 to 45 minutes. The test–retest reliability, and the concurrent and predictive validity of the instrument, have been well established.[80, 83, 84] The instrument has been used in several large studies of health service delivery such as the Comprehensive Community Mental Health Services for Children and Their Families[80, 85] and has been used by many states to determine eligibility for state funded programs. There are two versions of the instrument, one for children 5 to 19 years and another (PECFAS) for younger children ages 0 to 6. Advantages of the CAFAS are that it has: a Spanish version; dimensional subscales as well as a total score; severity level cut offs; good and ample psychometric properties; wide age range with two parallel versions; and parent and child versions. In addition, the CAFAS evaluates the home environment as a possible barrier to the child's functioning, and has been widely used in large surveys. Disadvantages include the need to be used in conjunction with another instrument or the longer version of the CAFAS interview (45 minutes) since it requires some knowledge of the child to derive the impairment ratings based on the 10 minute CAFAS instrument. The CAFAS also includes several psychiatric symptoms with the impairment items.

Functional Impairment Scale for Children and Adolescents (FISCA): The FISCA is based conceptually and structurally on the CAFAS and for this reason it will be described only in terms on how it differs from the CAFAS. The instrument has fewer items (85) than the CAFAS to measure impairment since it eliminated from the CAFAS some of the criteria for determining minimal impairment. Information is obtained only from parent reports. It is relatively knew, and it does not appear to have been used in any large survey or population based study.[69] The instrument has good construct validity for severely impaired hospitalized adolescents.[69] The

FISCA predicted hospital re-admissions among severe hospital patients in 50% of the cases.[69]

The instrument has the same disadvantages stated for the CAFAS, except for the fact that it is

relatively shorter. In addition, there is only a parent version, there is no younger child version,

and there is a need for further psychometric tests of the instrument particularly for younger and

less severely impaired populations.

Behavioral and Emotional Rating Scale (BERS). The 52 item BERS is a scale of a

child's social and emotional strengths and was originally designed for completion by an adult

who knew the child well.[86] More recently a youth rating scale of the BERS was developed and

tested by its authors with good test-retest reliability and convergent validity.[87] The adult and

youth BERS assess five dimensions of childhood strengths: interpersonal strengths, involvement

with family, intrapersonal strengths, school functioning and affective strengths. Subscales scores

for each of these domains based on national norms are provided together with a total score.

However, the scale does not have cut off severity scores. The scale has been mostly used for

placement in special education classes, child welfare and mental health settings to target goals

for individual treatment plans. Perhaps it is one of the scales with the most psychometric

properties. The adult informant BERS was standardized on a nationally representative sample of

children and adolescents and compared to children with behavioral and emotional difficulties.[88]

Data from this study was used to determine its content, and discriminant validity. Test-retest

reliability of the adult informant instrument has also been reported.[89,90] Its convergent and

construct validity have also been established.[91] One advantage of this scale is its wide age range,

since it can be administered with parents as informants from age 0 to 18. There is also an

adolescent version of the instrument. In addition, it has excellent psychometric properties for

both parent and child informants. The disadvantages of the instrument are related to the

following: the instrument needs to be administered to a person who knows the child well; there is no Spanish translation available; there are no severity cut off points; and testing of the instrument has been mostly limited to non-Latino white participants. Another disadvantage for a large survey administration is the need for various informants in order score the instrument.[92] The scale takes about 20 minutes to administer which may be too long for a large scale study.

Brief Impairment Scale (BIS): The BIS is a short multidimensional scale that assesses three domains of functioning; school/work, interpersonal relations and self-fulfillment.[93] The instrument has several advantages. It purposely excludes data on psychiatric symptomatology, is respondent based and scoring is independent of clinical judgment by the interviewer, in other words, it can be used by trained lay persons. The BIS has excellent psychometric properties. It has very good internal consistency and test-retest reliability in clinical and community populations of Puerto Rican children 4 to 17 years of age for the three subscales and total scores. There is an English and Spanish version of the instrument. The disadvantages include: no child version (only a parent version is available); not tested for children younger than 4; and no psychometric properties are available in an English speaking population. Although a cut off for children that are considered most impaired is provided, there are no cut offs by level of severity. The BIS has not been used in large surveys with the exception of the island of Puerto Rico survey[19] in which it was tested.

Child World Health Organization Disability Assessment Scale (C-WHO-DAS)

A new instrument to measure child disability, the C-WHO-DAS, has been developed that is based on the assumption that cultures and contexts vary and are important to consider in assessing functional disability. The C-WHO-DAS was adapted for children from the adult WHODAS II[94] by the DSM-5 Impairment/Disability Workgroup.[95] The adaptation process

27

included making sure that the items could be well understood by children and their families, and that the items were consonant with the basic assumptions of child disability described in the ICF-CY.[6]

In the C-WHO-DAS, functional disability is a separate construct from the disorder or illness and does not necessarily have to occur as a result of the same. The C-WHO-DAS has two versions, one to be completed by parents about their children ages 0 to 17 and another parallel form for youth 12 years and above to complete about themselves. In addition, there is a clinician version. Consonant with the ICF-CY, the three versions of the C-WHO-DAS assess the following domains: understanding and communicating, getting around (mobility), self-care, getting along with people, life activities (school and non-school) and participating in society. The time frame for all three versions is the last 30 days.

The parent and the youth versions begin with a global rating of overall health in the past 30 days, rated on a five-point scale (very good, good, moderate, bad, and very bad). This is followed by 34 items divided across the different domains. For each item the respondent considers the level of difficulty on a five point scale (none, mild, moderate, severe, extreme/cannot do). At the end of the questionnaire the participant is asked to provide an overall rating of how much his/her difficulties interfered with the child's life (using the same five point scale); the number of days (or the last thirty) the difficulties were present; the number of days the child was unable to carry out usual activities or the child had to cut back on usual activities.

The C-WHO-DAS has an English version and has been translated to Kinyarwanda for use among children in Rwanda. At present there is only one study that has tested the psychometric properties of this instrument and that is a study among children in Rwanda.[96] The C-WHO-DAS was used in the clinical trials of DSM 5. However, to this date there are no

publications on the psychometrics of this instrument in U.S. clinical samples or in any other sample of children besides the Rwanda study. In this study the English C-WHO-DAS was translated and adapted for use among children in Rwanda by a bilingual committee using state of the art methods.[96] The C-WHO-DAS was tested for internal consistency, test retest reliability, construct and divergent validity. The reliability of the instrument was tested with a sub-sample of 30 children who had participated in the overall survey of 367 children and their caregivers, 67% of whom had been referred by local health workers for a mental health problem. Results showed a very good internal consistency (Cronbach's alpha for all items, .84) for the whole sample. Test-retest reliability as obtained from the sub-sample was also very good (ICC=.83). The overall C-WHO-DAS scores were not significantly different for boys and girls, or by age, independent of who was reporting, child or parent. However, the correlation between the parent and child report scores was poor (r=.32). A confirmatory factor analysis of the scale showed a six factor model with the factors comprising the six domains of the C-WHO-DAS. Comparison of the C-WHO-DAS with the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI KID)[97] showed the C-WHO-DAS to be moderately correlated with major depressive episode and post-traumatic stress disorder diagnoses and less correlated with anxiety and conduct disorder. Mean C-WHO-DAS scores were significantly higher for those meeting criteria for any disorder in the MINI-KID as compared with children that did not meet disorder criteria.

The advantages of the C-WHO-DAS are: 1) it has a wide age range from 0 to 18 (the adult WHO-DAS can be used for those 19 to 22 years), 2) it has parent and child versions, 3) it separates psychiatric symptoms from impairment, 4) it is based on the ICF-CY, which is the internationally recognized classification of impairment, 5) it covers six domains consonant with the federal register definition of SED, 6) severity of impairment is determined by impairment

item and by domain, 7) it is respondent based, not requiring prior knowledge of the child. The main disadvantage is that there is no published psychometrics of the English version of the instrument, and no translation or psychometric tests for Spanish speaking populations.

Summary and Recommendations

From this review of the literature on impairment and the existent measures to assess the construct, *it is clear that all the scales reviewed differed in meeting the ideal 8 criteria we specified in this review, their adaptability to a large survey and our interpretation of the Federal Register SED definition. None of the scales met the ideal 8 criteria specified.* Several challenges were identified that made it difficult to choose a measure that would be ideal for the measurement of impairment. With the exception of the C-WHO-DAS, all of the impairment measures were developed prior to the development of the International Classification of Functioning for Youth (ICF-CY).[6] This lack of a classificatory system that would provide a standard definition of impairment was possibly the main reason why the instruments varied so much in their conceptualization of impairment, and as a consequence varied so much in the way of assessing the construct. The first instruments developed were mostly global measures that did not consider the importance of measuring the child's impairment according to the child's context (family, school, peer, community or culture). Although the scales recognized the importance of items that would tap into these domains, it was not possible to determine if the child was functioning well in one or two domains but functioning very poorly in the other domains, information that is necessary for prevention and treatment. However, the global scales, including the C-GAS, CIS, and the SDQ impact score, had the advantage of being very brief with no more than 3 to 5 minutes of administration, precisely because there were no sub-scales by domain. A

major disadvantage of two of these global scales (C-GAS, SDQ) is that the interviewer makes the judgment of the child's impairment and therefore requires prior knowledge of the child. Dimensional scales of impairment (CAFAS, SAICA, BERS, FISCA, BIS, C-WHO-DAS) are much more lengthy, and take more time to administer. Not all scales provided cut off scores by level of severity, or had parent and child versions, or were completely separate from psychiatric symptoms, or could be administered to a wide age range, or were respondent based without requiring prior knowledge of the child. However, all the scales, with the exception of the C-WHO-DAS had good psychometric properties in the U.S. population.

Perhaps the greatest challenge for the choice of a scale is related to what is considered substantial impairment, given the different criteria and methods used by the different instruments that provide severity cut offs. Furthermore, severity scores are determined for a wide age range of children without considering how severity might differ by developmental stage. Even if the choice is to consider only cut offs of severe impairment as part of what constitutes substantial impairment, the severity score of one instrument gives a different prevalence rate as compared to the severity score of another instrument. It is no wonder then, that prevalence rates of impairments differ so much from one study to another if different instruments are used. *There is a need to agree on a standard definition and method for estimating substantial impairment that is developmentally grounded and culturally sensitive.*

The other challenge in planning a national survey for the measurement of SED is the choice of the informant. The general consensus among researchers and clinicians is that both informants (parent and children) are necessary given that parents and children differ significantly in their impairment scores and that for certain behaviors children (12 years and above) are better informants, while for others, parents are better informants. The choice to make if both informants

are to be assessed, is how to best combine these two informants to arrive at a prevalence rate of impairment. An alternative would be to give the instrument to youth 12 to 22 years of age since they are considered adequate informants and to parents in relation to their children aged 0 to 11. The inclusion of two informants in a survey is costly and this alternative covers the entire age range with one informant for different ages. Having two informants (i.e. parent and child) for one age group (12 to 22) can inflate the rates in comparison to the younger age group (0 to 11) in which only the parent can provide the information. So if budgetary constraints limit the informant to one, then one choice is to include only the parent for ages 0 to 11 and the youth for ages 12 to 22.

Of all the instruments reviewed, only the CAFAS had two versions for different ages, one for children 0 to 6 years and another for those 7 to 18. Although the C-GAS and C-WHO-DAS potentially could be administered to children younger than 4 years, these instruments do not have published psychometric properties for children younger than 6 years of age. The C-GAS for example included children 6 to 17, and the testing of the C-WHO-DAS with children from Rwanda included children 10 to 17. Thus, it is not known how well these instruments work with children below the age of 6. For example, with preschoolers impairment would probably be better expressed through parental burden, as recommended by the expert consultants. It is logical then to think that for very young children or infants (i.e, 0 to 5 years) a different version of impairment instrument should be used. Unfortunately, only one very long measure, the CAFAS, has that possibility. *There is definitely a need for the development of impairment measures or versions of existent measures that are tailored to this young age group. The ideal measure would have three versions: one for infants and preschoolers, another for children 5 to 11 years of age, and still another for adolescents 12 to 18 years of age.*

As previously stated, at present none of the impairment measures meet the ideal criteria for complying with the Federal Register definition, the recommendations of the expert panel, and cost considerations for administration in a large survey. SAMHSA needs to decide whether the development of a new instrument that meets all the requirements is a worthwhile investment for the near future. Another option is to work on existing measures with secondary analyses of existing data sets. There are a number of alternatives that can be followed if SAMHSA has time to invest in further development of existent measures. These are my recommendations.

1. If there is no time and money to invest in the development of an existing measure, I recommend the use of the Columbia Impairment Scale. It seems the instrument with the best alternatives for a large scale survey. The CIS is short, has severity cut offs, is respondent based, has child and parent versions, and it has excellent psychometric properties in English and Spanish. Although it is a global scale and therefore cannot be scored by contextual domain, it contains items that refer to the child's functioning at school, and with family and peers. Multidimensional scales that can produce impairment scores by domain are much longer and a score by domain is more useful for prevention and intervention. If what is needed is just one impairment score to determine severity, similar to what is required for the definition of SED, then a score by domain is not necessary, even though it is the ideal standard. The main limitation of the CIS is that there is no version for infants and preschoolers. Different versions for these two developmental ages would need to be developed. If these two populations are not to be surveyed, then there is no need to do another study, given that there is recent national information on impairment rates assessed with the CIS and reported by Olfson et al.[51] However, if pre-school children are to be included in a

survey, then there would be a need to develop a pre-school CIS or develop a different measure for this age group.

2. If there is time for further development of an existing measure, I suggest that secondary analyses be done of the C-WHO-DAS using data from the American Psychiatric Association DSM 5 clinical trials in which the instrument was administered to approximately 1,000 U.S. children in treatment. This instrument has the advantage over all other instruments, since it is based on the ICF-CY[6] and therefore can be used internationally, it is relatively short (36 items), it conceptually fulfills all the necessary criteria for the measurement of impairment in children, and is well suited to the Federal definition of SED. The C-WHO-DAS is respondent based, has child, parent and clinician versions. The fact that the psychometric properties of the C-WHO-DAS have been shown to be very good when administered to children in Rwanda[96] is promising. The psychometric properties reported in this study were relatively good considering that the authors used the scoring system developed for the adult WHO-DAS. Scoring for children and adolescents needs to be developed. In addition, and if there is time, it would be of interest to further develop different versions of the instrument specifically tailored for infants and preschoolers. Permission for the use of this instrument from the WHO is easy to obtain, since several countries are using the instrument with the endorsement of the WHO. However, this recommendation is limited by the fact that the C-WHO-DAS was developed from an adult instrument. The question remains as to whether a new instrument should be developed for children based solely on the ICF-CY, given that

the present C-WHO-DAS was based on the adult WHO-DAS that is based on the adult international classification of functioning (ICF) and not the child classification.

3. The CAFAS impairment measure is interviewer based and therefore needs to be administered in conjunction with another measure of psychiatric symptomatology so that interviewers can make the impairment judgement. This is the reason why the impairment items take so long, otherwise they would take 10 minutes to administer by for example a clinician who knows the child well. However, this instrument has a preschool and infant version for children 0 to 5 years which is an advantage that no other instrument has. Further work could be done for the development of a respondent base impairment measure of the CAFAS. Alternatively, the impairment items of the CAFAS could be tested together with other short psychiatric symptom instruments such as the SDQ, or longer psychiatric interviews such as the preschool DISC or PAPA.

4. We know that it is much cheaper to obtain rates of impairment and SED when only one informant is used. But we also know that using both parent and child informants (and for some diagnoses such as ADHD, a teacher, as well) are important and necessary to obtain accurate estimates of both psychiatric disorder and impairment. The issue then is how to use and combine data from these informants in the most efficient and least costly way. As discussed above these issues are still not resolved in the literature and further work is necessary. Analyses of secondary data sets of epidemiology studies that contain parent and child informants as well as impairment measures could shed more light on this unsettled issue. Of importance would be to analyze how rates change when using two informants in one age range and one

informant in another age range, as well as how rates change when only one informant

such as the child is used for ages 12 and above and just parents for ages 0 to 11. It

would be of interest to know how the sensitivity and specificity of the rates change by

using different informants by age and by using different ways of combining the

information obtained by different informants. Dr. Prudence Fisher from the

Department of Psychiatry at Columbia University has more than 10 epidemiology

data sets from Puerto Rico and the United States that could be used for this purpose.

In summary, there is a need for research to focus on the development of more

developmentally and culturally appropriate measures for the measurement of impairment. Cost-

containment strategies will be emphasized and enforced by payers during the next decade, and

greater demand for children's impairment measures will come into effect.[95] When one considers

how much progress has been made in the field of psychiatric disorder measurement, it is not

unreasonable to hope that a similar investment of time and money could be equally fruitful in

producing a range of reliable, valid measures of impairment particularly now that we can count

on a classificatory international system of impairment such as the ICF-CY.[6]

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 3er ed., Author; Washington, DC: 1987.

2. US Department of Health and Human Services. Federal Register: No 58:29425, "*Substance Abuse and Mental Health Services Administration (notices),*" 1993.

3. American Psychiatric Association. *The Diagnostic and Statistical Manual of Mental Disorders*. Fifth Edition. Author; Arlington, VA, 2013.

4. World Health Organization. ICF-CY, *International Classification of Functioning, Disability and Health; Children and Youth Version*. Geneva: World Health Organization, 2007.

5. Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality. *Serious emotional disturbance (SED) expert panel meeting: Operationalizing the SED definition for the production of national and state prevalence estimates, September 8, 2014, Gaithersburg, MD* [meeting notes]. Rockville, MD: Author, in press.

6. Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality. *Serious emotional disturbance (SED) expert panel meeting: Instrumentation and measurement issues when estimating national and state prevalence of childhood SED, November 12, 2014, Gaithersburg, MD* [meeting notes]. Rockville, MD: Author, in press.

7. Bird HR, Gould MS. The use of diagnostic instruments and global measures of functioning in child psychiatry epidemiological studies. In F. C. Verhulst & M. Hans (Eds.), *The Epidemiology of Child and Adolescent Psychopathology* (pp. 309–330). New York: Oxford University Press, 1995.

8. Canino G, Costello EJ, Angold A. Assessing functional impairment and social adaptation for child mental health services research: A review of measures. *Ment Health Serv Res*. 1999;1(2):93-108.

9. Florian L, Hollenweger J, Simeonsson RJ, Wedell K, Riddell S, Terzi L, et al. Cross-cultural Perspectives on the Classification of Children with Disabilities, Part 1: Issues in the classification of children with disabilities. *J Special Edu*. 2006;40(1):36-45.

10. Simeonsson RJ, Scarborough AA, Hebbeler KM. ICF and ICD codes provide a standard language of disability in young children. *J Clin Epidemiol*. 2006;59(4):365-373.

11. Bird HR, Andrews H, Schwab-Stone M, Goodman S, Dulcan M, Richters J, et al. Global measures of impairment for epidemiologic and clinical use with children and adolescents. *Intern J Methods Psychiatric Res*. 1996;6:295-307.

12. Costello EJ, Egger HL, Angold A. The developmental epidemiology of anxiety disorders: phenomenology, prevalence, and comorbidity. *Child Adolesc Psychiatr Clin N Am*. 2005;14(4): 631-648, vii.

13. Ezpeleta L, Keeler G, Erkanli A, Costello EJ, Angold A. Epidemiology of psychiatric disability in childhood and adolescence. J Child Psychol Psychiatry. 2001;42(7):901-914.

14. Winters NC, Collett BR, Myers KM. Ten-year review of rating scales, VII: Scales assessing functional impairment. *J Am Acad Child Adolesc Psychiatry*. 2005;44(4):309-338; discussion 339-342.

15. Angold A, Erkanli A, Farmer EM, Fairbank JA, Burns BJ, Keeler G, et al. Psychiatric disorder, impairment, and service use in rural African American and white youth. *Arch Gen Psychiatry*. 2002;59(10):893-901.

16. Briggs-Gowan MJ, Horwitz SM, Schwab-Stone ME, Leventhal JM, Leaf PJ. Mental health in pediatric settings: distribution of disorders and factors related to service use. *J Am Acad Child Adolesc Psychiatry.* 2000;39(7):841-849.

17. Cuffe SP, Waller JL, Addy CL, McKeown RE, Jackson KL, Moloo J, et al. A longitudinal study of adolescent mental health service use. *J Behav Health Serv Res*. 2001;28(1):1-11.

18. Alegria M, Canino G, Lai S, Ramirez RR, Chavez L, Rusch D, et al. Understanding caregivers' help-seeking for Latino children's mental health care use. *Med Care*. 2004;42(5):447-455.

19. Canino G, Shrout PE, Rubio-Stipec M, Bird HR, Bravo M, Ramirez R, et al. The DSM-IV rates of child and adolescent disorders in Puerto Rico: prevalence, correlates, service use, and the effects of impairment. *Arch Gen Psychiatry*. 2004;61(1):85-93.

20. Lollar DJ, Simeonsson RJ. Diagnosis to function: classification for children and youths. *J Dev Behav Pediatr*. 2005;26(4):323-330.

21. Lehman, AF, Alexopoulos, GS, Goldman H, Dilipjeste MD, Ustun B. Mental disorders and disability: Time to reevaluate the relationship? In: *A Research Agenda for DSM-V*. Edited by David J. Kupfer, Michael B. First, and Darrel A. Regier. American Psychiatric Press Inc., pp. 201-218, 2002.

22. Goodman R, Ford T, Simmons H, Gatward R, Meltzer H. Using the strengths and difficulties questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *Br J Psychiatry*. 2000;177:534-539.

23. Wille N, Bettge S, Wittchen, H-U, Ravens-Sieberer U and the BELLA study group. How impaired are children and adolescents by mental health problems? Results of the BELLA study. *Eur Child Adolesc Psychiatry*. 2008;[Suppl 1] 17:42-51.

24. Collishaw S, Goodman R, Ford T, Rabe-Hesketh S, Pickles A. How far are associations between child, family and community factors and child psychopathology informant-specific and informant-general? *J Child Psychol Psychiatry*. 2009;50(5):571-580.

25. Feaster D, Brincks A, Robbins M, Szapocznik J. Multilevel models to identify contextual effects on individual group member outcomes: a family example. *Fam Process*. 2011;50(2):167-183. PMCID: 3736581.

26. Grills AE, Ollendick TH. Issues in parent-child agreement: the case of structured diagnostic interviews. *Clin Child Fam Psychol Rev*. 2002;5(1):57-83.

27. Herjanic B, Reich W. Development of a structured psychiatric interview for children: agreement between child and parent on individual symptoms. *J Abnorm Child Psychol*. 1997;25(1):21-31.

28. Hodges K, Gordon Y, Lennon MP. Parent-child agreement on symptoms assessed via a clinical research interview for children: the Child Assessment Schedule (CAS). *J Child Psychol Psychiatry*. 1990;31(3):427-436.

29. Mitsis EM, McKay KE, Schulz KP, Newcorn JH, Halperin JM. Parent-teacher concordance for DSM-IV attention-deficit/hyperactivity disorder in a clinic-referred sample. *J Am Acad Child Adolesc Psychiatry*. 2000;39(3):308-313.

30. Grills AE, Ollendick TH. Multiple informant agreement and the anxiety disorders interview schedule for parents and children. *J Am Acad Child Adolesc Psychiatry*. 2003;42(1):30-40.

31. Jensen P & Rubio-Stipec M, Canino G, Bird H, Dulcan M, Schwab-Stone M, Lahey B. Parent and child contributions to child psychiatric diagnosis: are both informants always necessary? *J Am Acad Child Adolesc Psychiatry*. 1999;38(12):1569-1579. PMID: 10596258.

32. Schwab-Stone ME, Shaffer D, Dulcan MK, Jensen PS, Fisher P, Bird HR, et al. Criterion validity of the NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3). *J Am Acad Child Adolesc Psychiatry*. 1996;35(7):878-888. PMID: 8768347

33. Achenbach TM. Diagnosis, assessment, and comorbidity in psychosocial treatment research. *J Abnorm Child Psychol*. 1995;23(1):45-65.

34. Merikangas KR, He JP, Burstein M, Swanson SA, Avenevoli S, Cui L, et al. Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication--Adolescent Supplement (NCS-A). *J Am Acad Child Adolesc Psychiatry*. 2010; 49(10):980-989. PMCID: 2946114.

35. Roberts RE, Alegria M, Roberts CR, Chen IG. Concordance of reports of mental health functioning by adolescents and their caregivers: a comparison of European, African and Latino Americans. *J Nerv Ment Dis*. 2005;193(8):528-534.

36. Bidaut-Russell M, Reich W, Cottler LB, Robins LN, Compton WM, Mattison RE. The Diagnostic Interview Schedule for Children (PC-DISC v.3.0): parents and adolescents suggest reasons for expecting discrepant answers. *J Abnorm Child Psychol*. 1995;23(5):641-659.

37. Ribera JC, Canino G, Rubio-Stipec M, Bravo M, Bauermeister JJ, Alegria M, et al. The Diagnostic Interview Schedule for Children (DISC 2.1) in Spanish: reliability in a Hispanic Population. *J Child Psychol Psychiat*, 1996;37(2):195-204. PMID: 8682899.

38. Waters E, Stewart-Brown S, Fitzpatrick R. Agreement between adolescent self-report and parent reports of health and well-being: results of an epidemiological study. *Child Care Health Dev.* 2003;29(6):501-509.

39. Comer JS, Kendall PC. A symptom-level examination of parent-child agreement in the diagnosis of anxious youths. *J Am Acad Child Adolesc Psychiatry*. 2004;43, 878-886.

40. Lewis AJ, Bertino MD, Bailey CM, Skewes J, Lubman DI, Toumbourou JW. Depression and suicidal behavior in adolescents: a multi-informant and multi-methods approach to diagnostic classification. *Front Psychol*. 2014;5:766. PMCID: 4101965.

41. Loeber R, Green SM, Lahey BB, Stouthamer-Loeber M. Differences and similarities between children, mothers, and teachers as informants on disruptive child behavior. *J Abnorm Child Psychol*. 1991;19(1):75-95.

42. Salbach-Andrae H, Klinkowski N, Lenz K, Lehmkuhl U. Agreement between youth-reported and parent-reported psychopathology in a referred sample. *Eur Child Adolesc Psychiatry*. 2009;18(3):136-143.

43. Frank SJ, Van Egeren LA, Fortier JL, Chase P. Structural, relative, and absolute agreement between parents' and adolescent inpatients' reports of adolescent functional impairment. *J Abnorm Child Psychol*. 2000;28(4):395-402.

44. Ablow JC, Measelle JR, Kraemer HC, Harrington, R. Luby, J, Smider, N, et al. The MacArthur three-city outcome study: Evaluating multi-informant measures of young children's symptomatology. *J Am Acad Child Adolesc Psychiatry*. 1999;38(12):1580-1590.

45. Carlston DL, Ogles BM. Age, gender, and ethnicity effects on parent-child discrepancy using identical item measures. *J Child & Fam Studies*. 2009;18(2):125-135.

46. van der Meer M, Dixon A, Rose D. Parent and child agreement on reports of problem behaviour obtained from a screening questionnaire, the SDQ. *Eur Child Adolesc Psychiatry*. 2008;17(8):491-7.

47. Klein JB, Lavigne JV, Seshadri R. Clinician-assigned and parent-report questionnaire-derived child psychiatric diagnoses: correlates and consequences of disagreement. *Am J Orthopsychiatry*. 2010;80(3):375-385.

48. Verhulst FC, van der Ende J. Agreement between parents' reports and adolescents' self-reports of problem behavior. *J Child Psychol Psychiatry*. 1992;33:1011-1023.

49. Rescorla LA, Bochicchio L, Achenbach TM, Ivanova MY, Almqvist F, Begovac I, et al. Parent-teacher agreement on children's problems in 21 societies. *J Clin Child Adolesc Psychol*. 2014;43(4):627-642.

50. Kramer TL, Phillips SD, Hargis MB, Miller TL, Burns BJ, Robbins JM. Disagreement between parent and adolescent reports of functional impairment. *J Child Psychol Psychiatry*. 2004;45(2):248-259.

51. Olfson M, Druss BG, Marcus SC. Trends in mental health care among children and adolescents. *N Engl J Med*. 2015;372(21):2029-2038.

52. Simpson GA, Bloom B, Cohen RA, Blumberg S, Bourdon KH. U.S. children with emotional and behavioral difficulties: data from the 2001, 2002, and 2003 National Health Interview Surveys. *Adv Data*. 2005;(360):1-13.

53. Rothen S, Vandeleur CL, Lustenberger Y, Jeanpretre N, Ayer E, Gamma F, et al. Parent-child agreement and prevalence estimates of diagnoses in childhood: direct interview versus family history method. *Int J Methods Psychiatr Res*. 2009;18(2):96-109.

54. Achenbach TM, Bird HR, Canino G, Phares V, Gould MS, Rubio-Stipec M. Epidemiological comparisons of Puerto Rican and U.S. mainland children: parent, teacher, and self-reports. *J Am Acad Child Adolesc Psychiatry*. 1990;29(1):84-93.

55. Bird HR, Gould MS, Staghezza B. Aggregating data from multiple informants in child psychiatry epidemiological research. *J Am Acad Child Adolesc Psychiatry*. 1992;31(1):78-85.

56. Rubio-Stipec M, Fitzmaurice G, Murphy J, Walker A. The use of multiple informants in identifying the risk factors of depressive and disruptive disorders--are they interchangeable? *Soc Psychiatry Psychiatr Epidemiol*. 2003;38(2):51-58.

57. Lahey BB, Flagg EW, Bird HR, Schwab-Stone ME, Canino G, Dulcan MK, et al. The NIMH Methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA) Study: background and methodology. *J Am Acad Child Adolesc Psychiatry*. 1996;35(7):855-864.

58. Shaffer D, Fisher P, Dulcan MK, Davies M, Piacentini J, Schwab-Stone ME, et al. The NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3): description, acceptability, prevalence rates, and performance in the MECA Study. Methods for the Epidemiology of Child and Adolescent Mental Disorders Study. *J Am Acad Child Adolesc Psychiatry.* 1996;35(7):865-877.

59. Costello EJ, Angold A, Burns BJ, Erkanli A, Stangl DK, Tweed DL. The Great Smoky Mountains Study of Youth. Functional impairment and serious emotional disturbance. *Arch Gen Psychiatry*. 1996;53(12):1137-1143.

60. Bird HR, Yager TJ, Staghezza B, Gould MS, Canino G, Rubio-Stipec M. Impairment in the epidemiological measurement of childhood psychopathology in the community. *J Am Acad Child Adolesc Psychiatry*. 1990;29(5):796-803.

61. Angold A, Costello EJ. A test-retest reliability study of child-reported psychiatric symptoms and diagnoses using the Child and Adolescent Psychiatric Assessment (CAPA-C). *Psychol Med*. 1995;25(4):755-762.

62. Angold A, Costello EJ, Farmer EM, Burns BJ, Erkanli A. Impaired but undiagnosed. *J Am Acad Child Adolesc Psychiatry*. 1999;38(2):129-137.

63. Bird HR, Davies M, Fisher P, Narrow WE, Jensen PS, Hoven C, et al. How specific is specific impairment? *J Am Acad Child Adolesc Psychiatry*. 2000;39(9):1182-1189.

64. Sparrow SS, Balla DA, Cicchetti DV. *Vineland Adaptive Behavior Scales: Interview Edition-Survey Form Manual*. Circle Press: American Guidance Service, 1984.

65. Bruininks RH, Woodcock RW, Weatherman RF, Hill BK. *Scales of Independent Behavior-Revised*. Itasca IL: Riverside Publishing, 1996.

66. Harrison PL, Oakland T. *Adaptive Behavior Assessment System–Second Edition*. San Antonio, TX: The Psychological Corporation, 2003.

67. Fallon T, Winters N, Pumariega AJ, et al: CALOCUS: comparative and face validity, in Scientific Proceedings of the 48th Annual Meeting of the American Academy of Child and Adolescent Psychiatry (AACAP), New Research Abstract P25. Washington, DC, AACAP, 2001.

68. John K, Gammon GD, Prusoff BA, Warner V. The Social Adjustment Inventory for Children and Adolescents (SAICA): testing of a new semistructured interview. *J Am Acad Child Adolesc Psychiatry*. 1987;26(6):898-911.

69. Frank SJ, Paul JS, Marks M, Van Egeren LA. Initial validation of the functional impairment scale for children and adolescents. *J Am Acad Child Adolesc Psychiatry*. 2000;39(10):1300-1308.

70. Shaffer D, Gould MS, Brasic J, Ambrosini P, Fisher P, Bird H, et al. A children's global assessment scale (CGAS). *Arch Gen Psychiatry*. 1983;40(11):1228-1231.

71. Goodman R. The strengths and difficulties questionnaire: a research note. *J Child Psychology & Psychiatry*. 1997;38;581-586.

72. Goodman R. The extended version of the strengths and difficulties questionnaire as a guide to child psychiatric caseness and consequent burden. *J Child Psychol Psychiat*. 1999;40(5):791-799.

73. Goodman A, Goodman R. Strengths and difficulties questionnaire as a dimensional measure of child mental health. *J Am Acad Child Adolesc Psychiatry*. 2009;48(4):400-403.

74. Goodman A, Lamping DL, Ploubidis GB. When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): data from British parents, teachers and children. *J Abnorm Child Psychol*. 2010;38(8):1179-1191.

75. Klasen H, Woerner W, Wolke D, Meyer R, Overmeyer S, Kaschnitz W, et al. Comparing the German versions of the Strengths and Difficulties Questionnaire (SDQ-Deu) and the Child Behavior Checklist. *Eur Child Adolesc Psychiatry*. 2000;9(4):271-276.

76. Mullick MS, Goodman R. Questionnaire screening for mental health problems in Bangladeshi children: a preliminary study. *Soc Psychiatry Psychiatr Epidemiol*. 2001;36(2):94-99.

77. Bourdon KH, Goodman R, Rae DS, Simpson G, Koretz DS. The Strengths and Difficulties Questionnaire: U.S. normative data and psychometric properties. *J Am Acad Child Adolesc Psychiatry*. 2005;44(6):557-564.

78. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry*. 2001;40(11):1337-1345.

79. Rowe R, Maughan B, Moran P, Ford T, Briskman J, Goodman R. The role of callous and unemotional traits in the diagnosis of conduct disorder. *J Child Psychol Psychiatry*. 2010;51(6):688-695.

80. Hodges K, Xue Y, Wotring J. Outcomes for children with problematic behavior in school and at home served by public mental health. *J Emotional Beh Dis*. 2004;12(2):109-119.

81. Hodges K. *Child and Adolescent Functional Assessment Scale* (2nd Edition, Revised). Ann Arbor: Functional Assessment Systems; 2000.

82. Hodges K, Wotring J. The role of monitoring outcomes in initiating implementation of evidence-based treatments at the state level. *Psychiatr Serv.* 2004;55(4):396-400.

83. Hodges K., Wong MM. Psychometric characteristics of a multidimensional measure to assess impairment: The Child and Adolescent Functional Assessment Scale. *J Child Fam Studies*. 1996;5:445-467.

84. Pernice F, Gust J, Hodges K. *A Structured Interview for Collecting Objective Outcome Data in Clinical Settings*. Symposium presented at the 10th Annual Research Conference: A System of Care for Children's Mental Health: Expanding the Research Base, Tampa, FL, February 1997.

85. Hodges K, Wong MM. Use of the Child and Adolescent Functional Assessment Scale to predict service utilization and cost. *J Ment Health Adm*. 1997;24(3):278-290.

86. Epstein MH, Sharma HM. *Behavioral and Emotional Rating Scale: A strength based approach to assessment.* Austin, TX: PRO-ED; 1998.

87. Epstein MH, Mooney P, Ryser G, Pierce CD. Validity and reliability of the behavioral and emotional rating scale (2nd edition): Youth Rating Scale. *Res Soc Work Practice*. 2004;14(5):358-367.

88. Epstein MH. The development and validation of a scale to assess the emotional and behavioral strengths of children and adolescents. *Remedial and Special Education*. 1999; 20:258-262.

89. Epstein MH, Hertzog MA, Reid R. The Behavioral and Emotional Rating Scale: Long-term test-retest reliability. *Beh Dis*. 2001;26(4):314-320.

90. Harniss MK, Epstein MH, Ryser G, Pearson N. The Behavioral and Emotional Rating Scale: Convergent Validity. *J Psychoeducational Assessment*. 1999;17:4-14.

91. Epstein MH, Nordness PD, Nelson JR, Hertzog M. (2002). Convergent validity of the Behavioral and Emotional Rating Scale with primary grade-level students. *Topics Early Childhood Special Edu*. 2002;22:114-121.

92. Friedman KA, Leone P, Friedman P. Strengths-based assessment of children with SED: Consistency or reporting by teachers and parents. *J Child Fam Studies*. 1999;8:169-180.

93. Bird HR, Canino G, Davies M, Ramírez R, Chavez L, Duarte C, Shen S. (2005). The Brief Impairment Scale (BIS): A multidimensional scale of functional impairment for children and adolescents. *J Am Acad Child Adolesc Psychiatry*. 2005;44(7): 699-707.

94. Von Korff M, Crane PK, Alonso J, Vilagut G, Angermeyer MC, Bruffaerts R, et al. Modified WHODAS-II provides valid measure of global disability but filter items increased skewness. *J Clin Epidemiol*. 2008;61(11):1132-43.

95. Canino G, Fisher P, Alegria M, Bird H. Assessing Child Impairment in Functioning in Different Contexts: Implications for Use of Services and the Classification of Psychiatric Disorders," *Open J Med Psychology*. 2013;2(1): 29-34.

96. Scorza P, Stevenson A, Canino G, Mushashi C, Kanyanganzi F, Munyanah M, et al. Validation of the "World Health Organization Disability Assessment Schedule for children, WHODAS-Child" in Rwanda. *PLoS One*. 2013;8(3):e57725.

97. Sheehan D, Shytle D, Milo K, Janavs J, Lecrubier Y. *MINI International Neuropsychiatric Interview for Children and Adolescents, English Version 6.0.* Tampa, FL: University of South Florida; 2009.