American Journal on Intellectual and Developmental Disabilities

Using a Delphi Process to Update the Nisonger Child Behavior Rating Form

--Manuscript Draft--

Manuscript Number:	AJIDD-D-20-00095R2
Article Type:	Research Report
Keywords:	problem behavior; psychopathology; intellectual disability; Delphi method; developmental disabilities.
Corresponding Author:	Kelsey B Shively, PhD Nationwide Children's Hospital Columbus, OH UNITED STATES
First Author:	Kelsey B Shively, PhD
Order of Authors:	Kelsey B Shively, PhD
	Marc J. Tasse, Ph.D.
Manuscript Region of Origin:	UNITED STATES
Abstract:	Children and adolescents with intellectual and developmental disabilities (IDD) experience higher rates of psychopathology and problem behavior compared to typically developing children (de Ruiter et al., 2007; Einfeld et al., 2011). The Nisonger Child Behavior Rating Form (NCBRF; Aman et al., 1996; Tassé et al., 1996) is a rating scale that was developed and normed to specifically screen psychopathology and problem behavior in children and adolescents with intellectual and developmental disabilities (IDD). The current study aimed to update the NCBRF to the current Diagnostic and Statistical Manual of Mental Disorders, 5 th edition (DSM-5; American Psychiatric Association, 2013) and focus its assessment to screen for the most prevalent childhood psychopathologies that affect children and adolescents with IDD. The authors re-aligned the existing items to fit within a DSM-5 framework, and then used the Delphi method with an expert panel of professionals in IDD to evaluate the NCBRF item pool. This entire revision process included revising existing items, deleting items, and formulating new items. We obtained a final item pool after three iterations. The Delphi process and resulting item pool are described in this paper.

Abstract

Children and adolescents with intellectual and developmental disabilities (IDD) experience higher rates of psychopathology and problem behavior compared to typically developing children (de Ruiter et al., 2007; Einfeld et al., 2011). The Nisonger Child Behavior Rating Form (NCBRF; Aman et al., 1996; Tassé et al., 1996) is a rating scale that was developed and normed to specifically screen psychopathology and problem behavior in children and adolescents with intellectual and developmental disabilities (IDD). The current study aimed to update the NCBRF to the current Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5; American Psychiatric Association, 2013) and focus its assessment to screen for the most prevalent childhood psychopathologies that affect children and adolescents with IDD. The authors re-aligned the existing items to fit within a DSM-5 framework, and then used the Delphi method with an expert panel of professionals in IDD to evaluate the NCBRF item pool. This entire revision process included revising existing items, deleting items, and formulating new items. We obtained a final item pool after three iterations. The Delphi process and resulting item pool are described in this paper.

Keywords: problem behavior, psychopathology, mental disorders, intellectual disability, Delphi method, developmental disabilities, rating scale

Using a Delphi Process to Update the Nisonger Child Behavior Rating Form

The Nisonger Child Behavior Rating Form (NCBRF; Aman et al., 1996) was created out of a need for instruments to assess psychopathology and problem behavior in children and adolescents with intellectual and developmental disabilities (IDD). It is widely accepted that children with IDD experience higher rates of behavioral and emotional disorders, and/or demonstrate more problem behaviors (de Ruiter et al., 2007; Einfeld et al., 2011; Emerson, 2003**;** Koskentausta et al, 2002; Kurzius-Spencer et al., 2018; Mayes et al., 2011; Simonoff et al., 2008; Strømme, & Diseth, 2000; Taanila et al., 2003; Totsika et al., 2011). Even so, there are fewer rating scales or other tools used to assess these concerns in this population when compared with typically developing peers. Some rating scales that were developed for children and adolescents with intellectual disabilities include the Aberrant Behavior Checklist (ABC; Aman et al., 1985a, 1985b), Developmental Behavioral Checklist (DBC; Einfeld & Tonge, 1992, 2002), Reiss Scales for Children's Dual Diagnosis (Reiss Scales; Reiss & Valenti-Hein, 1994), and Behavior Problem Inventory-01 (BPI-01; Rojahn et al., 2001).

Of these rating scales, only the Reiss Scales (Reiss & Valenti-Hein, 1994) contains items and subscales aimed at assessing DSM-oriented diagnostic categories (i.e., anxiety disorder, attention deficit, conduct disorder, depression, psychosis, autism, and somatoform behavior), but also has a mix of other subscales that describe specific symptom presentations (i.e., anger/self-control, poor self-esteem, and withdrawn/isolated), and 10 items to address rare behaviors. As far as these authors are aware, the Reiss Scales have not been psychometrically evaluated past their development stage. These scales report low internal reliability, which may be influenced by a low number of items per subscale (i.e., maximum of 5 items per subscale) (Reiss & Valenti-Hein, 1994).

The NCBRF was adapted from the Child Behavior Rating Form (CBRF; Edelbrock, 1985), a behavior rating form used for typically developing children, to suit the assessment needs of children with IDD. The NCBRF has teacher and parent forms that include both pro-social (i.e., "Positive Social") and problem behavior subscales (i.e., "Problem Behavior"). The NCBRF has been widely used in large clinical trials studying the effects of psychotropic medications on children with IDD (Aman et al., 2002; Croonenberghs et al., 2005; Findling et al., 2017; Pandina et al., 2007; Shea et al., 2004; Snyder et al., 2002) and has been translated into several languages (Mircea et al., 2010; Tassé et al., 2000). Psychometric properties of the NCBRF's parent form have been reviewed twice in English speaking samples with mixed results (Lecavalier et al., 2004; Norris & Lecavalier, 2011). Lecavalier et al. (2004) evaluated the factor structure of the NCBRF in a sample of 330 children and adolescents with autism spectrum disorders (ASD) and concluded that the results of their study provided validity evidence for the NCBRF to be used in this population (Lecavalier et al., 2004). Norris and Lecavalier (2011) used more robust factor analytic technics to evaluate the NCBRF in a sample of 399 children and adolescents recruited from special education classrooms and an outpatient behavior support clinic. Results indicated that the original item assignment for the Social Competence items showed good fit, but the Problem Behavior items did not perform well (Norris & Lecavalier, 2011).

Study Aims

As described above, there already exists a number of behavior rating scales (ABC, DBC, Reiss Scales, BPI) that measure different problem behaviors in children and adolescents with intellectual disability and/or ASD. The only available rating scale that has DSM-oriented subscales is the Reiss Scales, which has become somewhat outdated. As a co-author of the NCBRF, we aimed to improve upon the existing scale by updating the items and subscale

structure of the original "Problem Behavior" subscales to align them with current theory of psychopathology and problem behavior in children and adolescents with IDD as a precursor to evaluating the factor structure and psychometric properties of the updated rating scale. The revision process focused on aligning these subscales with major diagnostic categories to increase its clinical applicability, and therefore renaming the "Problem Behavior" subscales to "Psychiatric Disorders and Behaviors of Concern." For the purposes of this study, children with IDD was defined as children diagnosed with intellectual disability (ID), global developmental delay (GDD), and/or ASD. Einfeld et al. (2011) completed a rigorous systematic review of studies reporting prevalence estimates of mental disorders in children and adolescents with IDD. The most prevalent comorbid diagnoses in children and adolescents with IDD were attention-deficit/hyperactivity disorders (ADHD), anxiety disorders, depression or mood disorders, and conduct or oppositional defiant disorders. Criteria for the four most common diagnostic categories found by Einfeld et al. (2011) have similarities with subscales from the original NCBRF. In addition to these four diagnostic groups, a subscale for self-injury and aggression remained in the updated NCRBF, as this is commonly seen in children and adolescent with IDD (Kanne & Mazurek, 2011; MacLean et al., 2010; McClintock et al., 2003), even though it does not constitute its own diagnostic category within the Diagnostic and Statistical Manual, 5th Edition (DSM-5; APA, 2013). We are confident that the clinical utility of this revised structure will be enhanced beyond the utility of the original NCBRF subscales structure of: Conduct Problem, Insecure/Anxious, Hyperactive, Self-Injury/Stereotypic, Self-Isolated/Ritualistic, and Overly Sensitive (see Aman et al., 1996; Tassé et al., 1996). Having a revised NCBRF to more clinically-based scales will provide greater ease for clinicians to identify the presence of mental health problems in children and youth with IDD.

Materials and Methods

NCBRF

The NCBRF included parent and teacher forms that assess problem behavior among children and adolescents ages 3 to 16 with IDD. The NCBRF included a total of 76 items: 10 Positive Social items and 66 Problem Behavior items. Positive Social items were rated on a 4-point rating scale (0 = not true, 1 = somewhat or sometimes true, 2 = very or often true, 3 = completely or always true), and Problem Behavior items were rated on a 4-point frequency/severity scale (0 = did not occur or was not a problem, 1 = occurred occasionally or was a mild problem, 2 = occurred quite often or was a moderate problem, 3 = occurred a lot or was a severe problem). Both parents and teachers complete the same battery of items, however the scoring algorithms differed slightly in number and item placement within subscales. There was a total of two empirically derived subscales under the Positive Social domain, and six empirically derived subscales under the Problem Behavior domain. Positive Social subscales included Compliant/Calm and Adaptive Social. Problem Behavior subscales included Conduct Problem, Insecure/Anxious, Hyperactive, Self-Injury/Stereotypic, Self-Isolated/Ritualistic, and Overly Sensitive. The teacher form's sixth Problem Behavior subscale was named "Irritable," instead of "Overly Sensitive," to reflect the slightly different composition of items. Scoring of the NCBRF is based on age norms.

Procedure

The process of revising the NCBRF was two-fold: the authors completed work prior to the Delphi method (henceforward referred to pre-Delphi procedures), and then the expert panel participated in the Delphi method. During the pre-Delphi procedures, existing NCBRF items were reviewed by the authors and (a) kept for inclusion into the revised item pool, (b) set aside

for editing, or (c) removed from the item pool. Inclusion into one of these three categories was determined by fit into DSM-5 criteria for selected childhood disorders, as well as behavior problems. Items that fit poorly into one of the five defined subscales were then revised, and new items were created to fill identified gaps in diagnostic criteria. The expert panel then reviewed all new and edited items through the Delphi method.

Pre-Delphi Procedures

Item Fit with Diagnostic Criteria. Existing NCBRF items were cross-referenced with DSM-5 criteria for the four most prevalent psychopathology diagnoses seen in children and adolescents with IDD: ADHD (including primarily inattentive, primarily hyperactive/ impulsive, and combined subtypes), anxiety disorders (including generalized, separation, and specific phobias), depressive disorders, and ODD/CD. Items assessing self-injury and other severe problem behavior were also kept for use in a fifth subscale. Items were discarded if they did not have a clear link to the five identified subscales or if the main purpose of the item was to screen for diagnostic symptoms of ASD. Three items were deleted due to no clear link with the five identified subscales (i.e., “Exaggerates abilities or achievements,” “Overly excited, exuberant,” “Secretive, keeps things to self”). Seven items were discarded because they were related to diagnostic symptoms of ASD, and more specifically, restricted and repetitive behaviors and/or interests. The decision to remove items related to ASD was made because it was not a study aim to screen for symptoms associated with ASD, as there are many other rating scales exclusively assess ASD symptoms, including Autism Spectrum Rating Scale (ASRS; Goldstein & Naglieri, 2009), Childhood Autism Rating Scale, 2nd Edition (CARS-2; Schopler, Van Bourgondien, Wellman, & Love, 2010), Social Communication Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003), and Social Responsiveness Scale, 2nd edition (SRS-2; Constantino, J.N. & Gruber, C.P.),

to name a few. Additionally, ASD was considered part of the broader IDD population that was identified to serve as the normative group for this subscale.

Item Revision and Creation. After items were fit according to the DSM-5 symptomology, the authors examined the DSM-5 criteria list for unrepresented symptoms. As the nature of the rating scale is to "screen," and not to diagnose, the authors attempted to have items for at least 80% of symptoms, so that the disorder was adequately screened for. Additionally, not all diagnostic criteria require all symptoms to be met. For example, to meet criteria for ADHD, inattentive subtype, six of the listed nine symptoms must be present. The new and revised items were written in a structure that was similar to the original item structure of the NCBRF, with a focus on readability of items to keep them at reasonable reading level (i.e., equivalent of a 6th grade reading level or lower). Reading level of all items would then be assessed after the expert panel process.

Some literature on the revision of rating scales has suggested that during the revision phase, authors should create more items than will be needed for the final rating scale (DeVellis, 2017). In this revision of the NCBRF, however, the authors limited the total number of items so that the form could be completed in approximately 15 minutes or less, which is equated with ease of use. This decision was informed by the nature of the NCBRF target respondents including parents and caregivers of children with IDD seeking a psychological assessment due to developmental and/or behavioral concerns. There is an obvious tradeoff in scale development between length of the scale, its psychometric properties, and the burden on the respondent completing the assessment. Hence, the current study was mindful to avoid significant item additions that would increase the total time of administration of the NCBRF and resulting burden on the respondent. A maximum item limit was informed by other similar rating scales. The DBC

reported that their 96-item rating scale can be completed in 15-20 minutes (Dekker et al., 2002), and the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2000, 2001), a 120-item rating scale, can reportedly also be completed on average in 15 minutes. The authors decided that the new item battery would include no more than 100 total items, but would ideally be similar in number to the original NCBRF total item number, which was 76 items. This number of items in the item pool would be sufficient to thoroughly assess the target psychopathologies and behavioral categories, while avoiding a significant increase in burden on parent or caregiver respondents.

Expert Panel Members

Before inclusion in the final item pool, newly created and modified items were submitted to an expert panel for review. The use of an expert panel is a useful practice to assist in item writing and revisions (DeVellis, 2017). Potential experts were recruited from graduates of Ohio State University's Intellectual and Developmental Disabilities Psychology PhD program, psychologists from an assessment clinic within a large children's hospital, and professional colleagues of the two prior groups. A total of 10 experts in the field were invited to participate and 6 accepted. Three experts declined to participate due to their workload and one expert did not respond to the invitation.

The assembled panel included six experts. The aggregated professional duties of the expert panel members included extensive research and/or clinical practice across the lifespan of individuals with intellectual and developmental disabilities in a variety of settings and organizations. See Table 1 for information regarding panel members' degrees, years of experience, and clinical/research experience.

Insert Table 1 Here

Delphi Method

The panel process was guided by the Delphi method, a quasi-anonymous series of iterative rounds of consultation and gathering information that is designed to combine expert opinion into group consensus (Hsu & Sandford, 2007; Lynn, Laman, & Englebardt, 1999; McKenna, 1994). The Delphi method has been used in several different fields of study, most commonly employed in health and social sciences research. For a detailed review of the Delphi method and its iterations, the author refers the reader to Hsu and Sandford (2007). Potential expert panel members were provided with a brief overview of the study goals, copies of the NCBRF, and a detailed description of the Delphi method, including an anticipated time commitment and timeline for each round. Each questionnaire of the Delphi was sent to the panel members in a fillable portable document format (PDF) format via email. Panel members were given two weeks to respond to each round of questionnaires, which is the suggested best practice for the Delphi method (see Hsu & Sandford, 2007). The first author then took one week to compile results and prepare the next round. The first round of the Delphi method was sent to panel members in January of 2018 and the final round was concluded in March of 2018. The Delphi method on average requires three rounds to attain the criterion of success (Hasson, Keeney, & McKenna, 2000; Hsu & Sandford, 2007).

The first round of the Delphi method began with a document that introduced the aims of the study and gave an overview of the expert panel process, as well as a questionnaire. The process started with presenting the current NCBRF items with the diagnostic criteria for the corresponding target psychiatric disorders (i.e., major depressive disorder, anxiety disorders,

attention-deficit/hyperactivity disorder, oppositional defiant disorder, and conduct disorder). Panel members were then exposed to revised and newly created NCBRF items that tapped different diagnostic criteria that were not well represented by the original NCBRF item pool. There were 25 questions on the questionnaire sent to the expert panel members in the first round. Questions asked panel members to either: 1) decide between two similar items or indicate that both items should be used, 2) provide feedback on newly written items, or 3) indicate to which subscale an item should be assigned. Each question also provided panel members with the opportunity to write in a new item, edit the current item, or indicate if the presented item should be discarded. Additional overarching open-ended questions provided panel members with the opportunity to propose any other new item within specific subscales. The classic Delphi method implements only open-ended questions during round one, however, it is also common and acceptable to use more structured questions during the first round (Hsu & Sandford, 2007).

The purpose of the second and third round of the Delphi method was to present the panelists with the results of the previous round and gather additional quantitative feedback. Each question presented in round one was reported with the corresponding rates of responses in the round two questionnaire. If an item had a suggested edit from the previous round, the newly written item was presented in conjunction with the other responses from round one. Panelists were then asked to rank order the options for each item. Next, panel members were presented the cumulative rank ordering and asked if they agreed or disagreed with the cumulative ranking. Panel members that disagreed were asked to provide an explanation. The purpose of these iterative steps was to approach a majority agreement (i.e., more than 50% of panel members) among the panel members regarding the highest quality items.

Results

Item Review, Revision, and Creation

Psychiatric Disorders and Behaviors of Concern items were first reviewed for fit with diagnostic criteria for the following psychiatric disorders: ADHD (including primarily inattentive, primarily hyperactive/ impulsive, and combined subtypes), anxiety disorders (including generalized, separation, and specific phobias), depressive disorders, and oppositional defiant and conduct disorders. Items were also retained if they described self-injury and aggression. Based on this review, eight items were kept to screen for ADHD, 10 items were kept to screen for anxiety disorders, eight items were kept to screen for depressive disorders, and 11 items were kept to screen for oppositional defiant and conduct disorders. Four additional items were kept to screen for self-injury and aggression. See Table 1 for a list of the original Problem Behavior items re-assigned to diagnostic groups for subscale placement into the new Psychiatric Disorders and Behaviors of Concern subscales. Of the items kept and sorted into diagnostic groups, four were hypothesized to cross load on more than one diagnostic category based on fit with DSM-5 diagnostic symptoms. Seven of the original NCBRF items were removed because they primarily screened for ASD symptoms, which was not one of the psychiatric disorders retained during the updating of the NCBRF, and these items did not screen for any of the selected psychiatric disorders or self-injury and aggressive behavior. The deleted items included "Rocks body or head back and forth repetitively," "Has rituals such as head rolling or floor pacing," "Repeatedly flaps or waves hands, fingers, or objects (such as pieces of string)," "Repeats the same sound, word, or phrase over and over again," "Odd repetitive behaviors (e.g., stares, grimaces, rigid posture)," "Stubborn, has to do things own way," and "Engages in meaningless, repetitive body movements."

Insert Table 2 Here

After all the NCBRF Psychiatric Disorders and Behaviors of Concern items had been assigned to diagnostic symptoms for the identified psychiatric disorders, the coverage of items to symptoms was assessed. Some diagnostic symptoms had up to three items that potentially screened for its presence, while other symptoms were left unassessed. In cases where a symptom had two or more items, items were flagged for expert panel review to decide which items better assessed for that specific diagnostic symptom. For symptoms that did not have any items to screen for its presence, items were written by the authors. As previously stated, the goal of the update to the NCBRF is to screen for high prevalence psychiatric disorders that appear in childhood and adolescence. The goal of item creation was to achieve a minimum of 80% coverage of these psychiatric disorder's symptoms.

The ADHD subscale had eight existing NCBRF items that matched well with diagnostic criteria. Three new items were written, and one existing item was edited to better screen for inattentive and hyperactive/impulsive symptoms included in the diagnostic criteria. ADHD diagnostic criteria have many symptoms listed, but to meet criteria, only a specific number of them are required to be present. For both ADHD subtypes (i.e., inattentive, and hyperactive/impulsive), the DSM-5 lists nine symptoms and requires that six symptoms (or more) must be present to meet criteria for that subtype. Existing NCBRF items covered 11 of the total 18 symptoms. One existing NCBRF item was edited ("Shifts rapidly from topic to topic"), and three new items were written ("Misplaces or loses things," "Forgetful in routine activities," and "Has trouble waiting his/her turn (i.e., in conversation, standing in line)") to better cover

ADHD symptoms. Four questions, each regarding a different item, were presented to the panel members for review regarding ADHD-related criteria.

For the anxiety disorders subscale, Generalized Anxiety Disorder (GAD) and Social Anxiety Disorder (SAD) criteria were matched with NCBRF items. Three new items were written, a previously written item was assigned to the anxiety category, and two pairs of items were flagged for expert panel review. The first new item written to screen for GAD symptoms related to sleep ("Gets tired easily"). Two other new items were written for the SAD symptom of, "The social situations almost always provoke fear or anxiety," which also notes that for children this may be, "expressed by crying, tantrums, freezing, clinging, shrinking, or failing to speak in social situations" (APA, 2013). These newly written items included, "Clings to adults," and "Refuses to talk in social situations." A previously written item for depressive disorders, "Sleep problems," was also assigned to the anxiety disorders category to screen for the GAD criteria, "Sleep disturbances (difficulty falling or staying asleep, or restless, unsatisfying sleep)" (APA, 2013). Existing NCBRF items ("Shy around others; bashful" and "Shy or timid behavior") possibly matched with the SAD symptom, "Marked fear or anxiety about one or more social situations in which the individual is exposed to possibly scrutiny by others" (APA, 2013) but were flagged for expert panel review due to their inexact match. The second pair of items included for panel review were, "Isolates self from others" and "Withdrawn, uninvolved with others." These items potentially screened for the SAD symptom of, "The social situations are avoided or endured with intense fear or anxiety" (APA, 2013). Due to the two items similarities and inexact match with criteria, they were presented to the panel members for review. Lastly, in addition to GAD and SAD criteria, a fourth item was written to screen for specific phobias ("Disproportional fear to objects or situations (i.e., bees, heights)"). A total of seven questions

regarding anxiety-related items (one question was placed in the depressive disorders category due to the item's intended cross loading) covering nine items were presented to panel members for feedback and review.

To screen for depressive disorders, two symptoms had three existing NCBRF items that potentially matched, and two additional symptoms were not covered in existing items. The depression criterion of, "Depressed most of the day, nearly every day as indicated by subjective report (e.g., feels sad, empty, hopeless) or observation made by others (e.g., appears tearful)" (which also includes a note about irritable mood in children qualifying for this criterion; APA, 2013), matched with three existing NCBRF items. These items included, "Crying, tearful episodes," "Unhappy or sad," and "Irritable." Due to similarities between the first two items, the authors decided to include these items in the expert panel review process. Panel members were asked if one item screened for the selected symptom better, or if both items should be kept. A second depression symptom, "Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated by either subjective account or observation)," also had three potential matching NCBRF items. These included, "Apathetic or unmotivated," "Isolates self from others," and "Withdrawn, uninvolved with others." Due to the similarities between the latter two items, it was decided that the expert panel members would provide feedback on which items were a better match, or if both items should be kept. Two other depression symptoms were not covered in existing NCBRF items for which the authors wrote new items for. The two newly written items included, "Unexpected change in appetite or weight," and "Sleep problems," to screen for their respective depressive disorder symptoms. A total of four questions covering six different items were presented to the panel members for feedback and review regarding depression-related items.

With regards to oppositional defiant and conduct disorders, one item was edited, another item was newly created, and three pairs of items were identified as potential matches to symptoms. The existing item, "Runs away from adults, teachers, or other authority figures," was simplified to "Runs away from home or school." One new item was written to account for the CD symptom of cruelty to animals ("Cruel or mean to animals"). Three symptoms in ODD and CD had two or more identified existing NCBRF items that potentially matched. The ODD symptom of, "Often argues with authority figures or, for children and adolescents, with adults" (APA, 2013), potentially matched with the items, "Argues with parents, teachers, or other adults," and "Talks back to teachers, parents, or other adults." The ODD symptom of, "Often actively defies or refuses to comply with requests from authority figures or with rules" (APA, 2013), similarly had two potential matches with the items, "Defiant, challenges adult authority," and "Disobedient." Lastly, the CD criterion of, "Often bullies, threatens, or intimidates others" (APA, 2013), potentially matched with the items, "Argues with other children or peers," and "Threatens people." It was decided that for these items, expert panel members would provide feedback as to which items better fit, or if a new item should be. Five questions covering eight total items were presented to panel members regarding ODD and CD items.

The last subscale included items for self-injury and aggression. Six existing NCBRF items were identified as fitting within this category. One new item was written, "Hits head on walls or doors." Panel members were asked to write additional Self-Injury and Aggression items that they believed should be added to this subscale. Three questions, including two open-ended questions, were presented to panel members regarding self-injury and aggressive behavior.

After items were assigned to respective diagnostic categories, there were three items that did not clearly match within a specific diagnostic category. These items included, "Overly

sensitive; feelings easily hurt," "Sudden changes in mood," and "Sulks, is silent and moody." Panel members were asked to assign these items to one of the four psychiatric disorders or indicate if the items should be discarded.

Delphi Method

Five of the six panel members responded for each of the first and second rounds. The panel member who did not respond in each round was different between these two rounds. All panel members responded to the final round. Results of each round of the Delphi method are presented in Tables 3 through 6, and are grouped by question type. Questions instructed panel members to 1) choose between 2 similar items, indicate desire to use both items, or discard both and write in a new item (see Table 3), 2) review newly written or edited items (see Table 4), 3) align an item with a subscale and then review the item (see Table 5), and lastly, 4) write in a new item (see Table 6) for the subscale "Self-Injury and Aggression."

For the first question type (Table 3), panel members were presented with two similar NCBRF items and asked to determine which item would be the best fit with the selected criteria for the respective diagnosis. Panel members were asked to indicate one of four choices: chose item 1, chose item 2, chose both items, or discard both items while proposing a new item. For one of the questions (i.e., "Argues with other children or peers" and "Threatens people"), two panel members indicated two responses (i.e., one panel members indicated an item, but also wrote in a new item, and a second panel member indicated that both items should be used, but also wrote in a new item). In round 2, panel members were then asked to rank order all of the items (including the use of both original items) from the results of round 1. In round 3, panel members were then asked if they agreed or disagreed with the top ranking item (or combination of items).

Question type 2 (Table 4) presented panel members with newly written or edited items along with the diagnostic criteria it was written to screen for. In round 1, panel members were asked to indicate if the item was acceptable as it was written, or if it needed editing. If panel members indicated that it needed editing, they were asked to write in their proposed edits. Round 2 presented panel members with all of the results of round 1, and then asked panel members to rank their preferred item. Lastly, in round 3, panel members were asked if they agreed or disagreed with the top ranking item.

Question type 3 (Table 5) presented panel members with items that were not already assigned to a subscale and asked them if the item fit within one of the four diagnostic subscales (i.e., ADHD, Anxiety, Depression, or ODD/CD), or if it should be discarded. Question type 4 (Table 6) was an open ended question that allowed panel members to write a new item for any behaviors they felt were not already represented in the self-injury and aggression subscale. For both question types 3 and 4, round 2 presented panel members with the results, and then asked if the item(s) needed editing. The last round asked panel members to rank their preferred item.

Insert Tables 3, 4, 5 & 6 Here

Throughout the Delphi method, there were two questions where panel members demonstrated an even split. For one item, panelists were split between two iterations of the same item (i.e., “Sulks, is silent or moody,” and “Is silent or moody, mopes”). Both items were included in the item pool for psychometric evaluation to help decide which item to retain in the final NCBRF-2. The panel members were also split between another set of items to screen for the oppositional defiant disorder symptom, “Often bullies, threatens, or intimidates others.” The

panel members were split between using the item “threatens people” alone, or using two items, “threatens people” and “bullies other.” Both items were used in the final NCBRF-2 item pool, as this decision captured the majority preference for, “threatens people” to be used, with the addition of another item that half of the panel members felt was important to include. The inclusion of these two sets of items did not bring the total number of items over the pre-determined maximum item limit (i.e., 100), and final inclusion will be informed by psychometric evaluation.

Results of the Delphi method provided feedback from panel members to include a total of 21 new items. Of those 21 new items, 6 items were accepted by the panel members as they were written by the authors, and 15 items were created or edited by panel members. See Table 3 for information regarding types of items in the final item pool. A list of the new or edited items that were reviewed and/or edited by the expert panel are presented in Table 7.

Insert Tables 7 & 8 Here

The iterative process of the Delphi method successfully used the experts’ feedback to arrive at a final item pool of 64 items to screen for the selected psychiatric disorders and behaviors of concern after three rounds. Consensus was defined as a majority of panel member’s responses (i.e., more than 50%), but average panel member agreement was 84%, with 19 of 26 questions obtaining 80% agreement or more. Together with the 10 positive social items, the NCBRF-2 item pool contained a total of 74 items. Review of the reading level of the resulting 74 NCBRF-2 items, using Flesch-Kincaid method, indicated an overall reading grade level of 6.4. A

reading level equivalent of 6th grade or below was targeted to accommodate the expected diversity in parental education level. A reading level of 6.4 grade level is comparable to the DBC's reading grade level of 7.3 (Einfeld & Tonge, 1995).

At the completion of the Delphi method, the final item pool was assembled. Items were formatted into a document that closely resembled the original structure of the NCBRF. Section II was renamed "Psychiatric Disorders and Behaviors of Concern" to reflect the modified content of the subscales.

Discussion

The aim of this study was to increase the clinical applicability of the original NCBRF Problem Behavior items/subscales. The original NCBRF, developed 25 years ago, consisted of six empirically derived Problem Behavior subscales that included some indirect overlap with criteria for psychiatric disorders: Conduct Problem, Insecure/Anxious, Hyperactive, Self-Injury/Stereotypic, Self-Isolated/Ritualistic, and Overly Sensitive (or "Irritable" in the teacher form). To focus and improve the clinical utility of the NCBRF, the items and subscale structure were revised to better reflect these childhood psychiatric disorders, with the addition of a fifth subscale for self-injury and severe aggressive behaviors, due to the high prevalence of these behaviors in this population (Kanne & Mazurek, 2011; MacLean et al., 2010; McClintock et al., 2003). Thus, the final NCBRF-2 Psychiatric Disorders and Behaviors of Concern section included five subscales: ADHD, Anxiety Disorders, Depressive Disorders, Oppositional Defiant and Conduct Disorders, and Self-Injury and Aggression. This shift of subscale focus from empirical descriptions of items to linkage with DSM-5 categories is an important step in recognizing mental health concerns in children and adolescents with IDD.

While the Reiss Scales is an available tool to screen for DSM-oriented diagnostic categories, its subscale structure presents psychometric limitations. Even though Reiss and Valenti-Hein (1994) concluded that the total score is a valid and reliable indicator of dual diagnosis, the brevity of individual subscales (5 items in each of 10 subscales) resulted in some subscales demonstrating low internal consistency. Additionally, this screening tool has not been psychometrically evaluated since it's development almost 30 years ago. The updated NCBRF fills the need for a current screening tool that can adequately screen for the most prevalent comorbid psychiatric childhood disorders in children and adolescents with IDD.

Identifying underlying psychopathology and/or high-risk problem behavior (i.e., self-injury and aggression) can lead clinical practitioners to making recommendations for interventions most appropriate for their presenting problem, instead of falling into diagnostic overshadowing. Additionally, researchers would benefit from being able to quantitatively measure symptoms of specific diagnostic categories to measure response to interventions that were created to alleviate those symptoms.

The Delphi method utilized a panel of six independent IDD experts with clinical experience in psychopathology to review the Psychiatric Disorders and Behaviors of Concern items and subscale placement. Previous scale development in the IDD field has used fewer experts or only used the authors of the scale for item creation. For example, development of the ABC relied on the four original authors to compile descriptions of behavior (Aman et al., 1985a), DBC development relied on the two original authors to assemble symptoms from case files and discriminate which behaviors were evidence of emotional or behavioral disorders (Einfeld & Tonge, 1995), and the Reiss Scales used four experts (i.e., two psychiatrists and two psychologists with expertise in IDD) to make comments on item revisions (Reiss & Valenti-

Hein, 1994). The Delphi method ensured that the items were reviewed objectively by a range of independent experts in the field of IDD. The Delphi method was completed in three rounds, and gathered feedback about item inclusion, newly created items, revised items, and item placement within subscales. The Delphi method created an item review process that fostered the combination of individual clinical judgement and cumulative group feedback to ensure items were rigorously reviewed through multiple steps of evaluation. Ultimately, the Delphi method resulted in a richer process of item development and review than if the original authors completed this work on their own, and it is recommended that this process be used in the future of scale development in the IDD field.

Limitations

The results of this study should be considered within a number of limitations. Although there is no consensus as to the actual recommended number of the experts to seat on a Delphi panel (Steurer, 2011), an expert panel size of 10 or more members has been cited as an ideal number of panel members to aim for in a Delphi method (Delbecq, Van de Ven, & Gustafson, 1975; Okoli & Pawlowski, 2004; McMillan et al., 2016). A limitation of our Delphi method is the small number of experts on our Delphi panel. Having said that, we do steadfastly argue that our six experts were highly qualified, experienced, and knowledgeable of the clinical and measurement issue at hand, which is ultimately the most important factor in reviewing an expert panel (Okoli & Pawlowski, 2004). In addition to the number of experts, the Delphi method may have benefited from recruitment occurring on a broader platform and through various mediums, such as through national organizations and affiliations. Another limitation of our methodology was our decision to set our consensus criterion among panel experts at >50%. Although there is no universally agreed-upon minimum proportion of experts defining “consensus” (Hasson et al,

2000) and although some researchers (see McKenna, 1994; Loughlin & Moore, 1979) have recommended a criterion of >50%, we concede that 50% may have led to premature decisions regarding some item revisions. Lastly, the authors completed pre-Delphi work, including reviews, edits, and creation of new items. This process was done intentionally to lighten the burden of work for the experts; however, the authors acknowledge that use of focus groups or expert panel meetings prior to employing the Delphi method may have enriched the data collection from experts.

Conclusion

More work is needed to continue the work to revise the NCBRF-2. The next steps include administering the NCBRF-2 to a large sample of participants to collect the necessary data to conduct psychometric analyses on these items and validate their ability to measure what they purport measuring. Nonetheless, we believe that the NCBRF-2 can become a useful assessment tool for clinicians and researchers needing an assessment measure using the current DSM framework of psychopathology to use with children and adolescents with IDD.

References

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA Preschool Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

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

Aman, M. G., De Smedt, G., Derivan, A., Lyons, B., Findling, R. L., & Risperidone Disruptive Behavior Study Group. (2002). Double-blind, placebo-controlled study of risperidone for the treatment of disruptive behaviors in children with subaverage intelligence. *American Journal of Psychiatry*, *159*(8), 1337–1346.

Aman, M. G., Singh, N. N., Stewart, A. W., & Field, C. J. (1985a). The Aberrant Behavior Checklist: A behavior rating scale for the assessment of treatment effects. *American Journal of Mental Deficiency*, *89*, 485–491.

Aman, M. G., Singh, N. N., Stewart, A. W., & Field, C. J. (1985b). Psychometric Characteristics of the Aberrant Behavior Checklist. *American Journal of Mental Deficiency*, *89*(5), 492–502.

Aman, M. G., Tassé, M. J., Rojahn, J., & Hammer, D. (1996) The Nisonger CBRF: A child behavior rating form for children with developmental disabilities. *Research in Developmental Disabilities, 17*(1), 41-57.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. Washington, DC: American Psychiatric Pub.

Constantino, J.N. & Gruber, C.P. (2012). Social Responsiveness Scale, Second Edition (SRS-2). Torrance, CA: Western Psychological Services.

Croonenberghs, J., Fegert, J. M., Findling, R. L., De Smedt, G., & Van Dongen, S. (2005). Risperidone in Children With Disruptive Behavior Disorders and Subaverage Intelligence: A 1-Year, Open-Label Study of 504 Patients. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*(1), 64–72. https://doi.org/10.1097/01.chi.0000145805.24274.09

de Ruiter, K. P., Dekker, M. C., Verhulst, F. C., & Koot, H. M. (2007). Developmental course of psychopathology in youths with and without intellectual disabilities. *Journal of Child Psychology and Psychiatry*, *48*(5), 498–507. https://doi.org/10.1111/j.1469-7610.2006.01712.x

Dekker, M. C., Nunn, R., & Koot, H. M. (2002). Psychometric properties of the revised Developmental Behaviour Checklist scales in Dutch children with intellectual disability. *Journal of Intellectual Disability Research*, *46*(1), 61–75.

Delbecq, A. L., Van de Ven, A. H., & Gustafson, D. H. (1975). *Group techniques for program planning: A guide to nominal group and Delphi processes.* Glenview, IL: Scott, Foresman and Company.

DeVellis, R. F. (2017). *Scale development: theory and applications* (4th ed.). Los Angeles: SAGE.

Edelbrock, C. S. (1985). Child behavior rating form. *Psychopharmacology Bulletin, 21*, 835-837.

Einfeld, S. L., Ellis, L. A., & Emerson, E. (2011). Comorbidity of intellectual disability and mental disorder in children and adolescents: A systematic review. *Journal of Intellectual*

and Developmental Disability, *36*(2), 137–143. https://doi.org/10.1080/13668250.2011.572548

Einfeld, S. L., & Tonge, B. J. (1992). Manual for the Developmental Behaviour Checklist. Clayton, Melbourne and Sydney: Monash University for Developmental Psychiatry and School of Psychiatry, University of New South Wales.

Einfeld, S. L., & Tonge, B. J. (1995). The Developmental Behavior Checklist: The Development and Validation of an Instrument to Assess Behavioral and Emotional Disturbance in Children and Adolescents with Mental Retardation. *Journal of Autism and Developmental Disorders*, *25*(2), 81–104.

Einfeld, S. L., & Tonge, B. J. (2002). Manual for the Developmental Behaviour Checklist: Primary Carer Version (DBC-P) & Teacher Version (DBC-T) (2nd. ed.). Clayton, Melbourne: Monash UniversityCentre for Developmental Psychiatry and Psychology

Emerson, E. (2003). Prevalence of psychiatric disorders in children and adolescents with and without intellectual disability. *Journal of Intellectual Disability Research*, *47*(1), 51–58.

Findling, R. L., Townsend, L., Brown, N. V., Arnold, L. E., Gadow, K. D., Kolko, D. J., … Aman, M. G. (2017). The Treatment of Severe Childhood Aggression Study: 12 Weeks of Extended, Blinded Treatment in Clinical Responders. *Journal of Child and Adolescent Psychopharmacology*, *27*(1), 52–65. https://doi.org/10.1089/cap.2016.0081

Goldstein, S., & Naglieri, J. A. (2009). *Autism spectrum rating scales (ASRS)*. Multi-Health System.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, *32*(4), 1008–1015.

Hsu, C.-C., & Sandford, B. A. (2007). The Delphi technique: making sense of consensus. *Practical Assessment, Research & Evaluation*, *12*(10), 1–8.

Kanne, S. M., & Mazurek, M. O. (2011). Aggression in children and adolescents with ASD: Prevalence and risk factors. Journal Of Autism And Developmental Disorders, 41(7), 926–937. https://doi.org/10.1007/s10803-010-1118-4

Koskentausta, T., Iivanainen, M., & Almqvist, F. (2002). Psychiatric disorders in children with intellectual disability. *Nordic Journal of Psychiatry*, *56*(2), 126–131. https://doi.org/10.1080/080394802753617944

Kurzius-Spencer, M., Pettygrove, S., Christensen, D., Pedersen, A. L., Cunniff, C., Meaney, F. J., Soke, G. N., Harrington, R. A., Durkin, M., & Rice, S. (2018). Behavioral problems in children with autism spectrum disorder with and without co-occurring intellectual disability. *Research in Autism Spectrum Disorders*, *56*, 61–71. https://doi.org/10.1016/j.rasd.2018.09.002

Lecavalier, L., Aman, M.G., Hammer, D., Stoica, W., & Mathews, G.L. (2004). Factor analysis of the Nisonger Child Behavior Rating Form in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *34*(6), 709–721.

Loughlin K. & Moore L. (1979). Using Delphi to achieve congruent objectives and activities in a pediatrics department. *Journal of Medical Education 54*, 101-106.

Lynn M. R., Laman E. L. & Englebardt S. P. (1998). Nursing administration research priorities: A national Delphi study. *Journal of Nursing Administration 28*(5), 7-11.

MacLean, W. E., Tervo, R. C., Hoch, J., Tervo, M., & Symons, F. J. (2010). Self-Injury among a Community Cohort of Young Children at Risk for Intellectual and Developmental

Disabilities. *The Journal of Pediatrics*, *157*(6), 979–983. https://doi.org/10.1016/j.jpeds.2010.05.052

Mayes, S. D., Calhoun, S. L., Murray, M. A., & Smith, L. A. (2011). Anxiety, depression, and irritability in children with autism relative to other neuropsychiatric disorders and typical development. *Research in Autism Spectrum Disorders*, *5*(1), 474–485. https://doi.org/10.1016/j.rasd.2010.06.012

McClintock K, Hall S, & Oliver C. (2003). Risk markers associated with challenging behaviours in people with intellectual disabilities: A meta-analytic study. Journal of Intellectual Disability Research, 47(6), 405–416.

McKenna H.P. (1994) The Delphi technique: a worthwhile approach for nursing? *Journal of Advanced Nursing 19*, 1221-1225.

McMillan, S. S., King, M., & Tully, M. P. (2016). How to use the nominal group and Delphi Techniques. International journal of clinical pharmacy, 38(3), 655-662.

Mircea, C. E., Rojahn, J., & Esbensen, A. J. (2010). Psychometric Evaluation of Romanian Translations of the Behavior Problems Inventory-01 and the Nisonger Child Behavior Rating Form. *Journal of Mental Health Research in Intellectual Disabilities*, *3*(1), 51–65. https://doi.org/10.1080/19315860903520515

Norris, M., & Lecavalier, L. (2011). Evaluating the validity of the Nisonger Child Behavior Rating Form – Parent Version. *Research in Developmental Disabilities*, *32*(6), 2894–2900. https://doi.org/10.1016/j.ridd.2011.05.015

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. Information & management, 42(1), 15-29.

Pandina, G. J., Bilder, R., Harvey, P. D., Keefe, R. S. E., Aman, M. G., & Gharabawi, G. (2007). Risperidone and Cognitive Function in Children With Disruptive Behavior Disorders. *Biological Psychiatry*, *62*(3), 226–234. https://doi.org/10.1016/j.biopsych.2006.09.036

Reiss, S., & Valenti-Hein, D. (1994). Development of a psychopathology rating scale for children with mental retardation. Journal of Consulting and Clinical Psychology, 62(1), 28–33. https://doi.org/10.1037/0022-006X.62.1.28

Rojahn, J., Matson, J. L., Lott, D., Esbensen, A. J., & Smalls, Y. (2001). The Behavior Problems Inventory: An instrument for the assessment of self-injury, stereotyped behavior, and aggression/destruction in individuals with developmental disabilities. *Journal of Autism and Developmental Disorders*, *31*(6), 577–588.

Rutter M, Bailey A & Lord C (2003). The Social Communication Questionnaire. Los Angeles: Western Psychological Services

Schopler, E., Van Bourgondien, M. E., Wellman, J., & Love, S. (2010). Childhood autism rating scale-second edition (CARS2): manual. Los Angeles: Western Psychological Services.

Shea, S., Turgay, A., Carroll, A., Schulz, M., Orlik, H., Smith, I., & Dunbar, F. (2004). Risperidone in the Treatment of Disruptive Behavioral Symptoms in Children With Autistic and Other Pervasive Developmental Disorders. *PEDIATRICS*, *114*(5), e634–e641. https://doi.org/10.1542/peds.2003-0264-F

Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric Disorders in Children With Autism Spectrum Disorders: Prevalence, Comorbidity, and Associated Factors in a Population-Derived Sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(8), 921–929. https://doi.org/10.1097/CHI.0b013e318179964f

Snyder, R., Turgay, A., Aman, M., Binder, C., Fisman, S., & Carroll, A. (2002). Effects of risperidone on conduct and disruptive behavior disorders in children with subaverage IQs. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(9), 1026–1036.

Steurer, J (2011). The Delphi method: an efficient procedure to generate knowledge. Skeletal Radiolology, 40, 959–961. DOI 10.1007/s00256-011-1145-z

Strømme, P., & Diseth, T. H. (2000). Prevalence of psychiatric diagnoses in children with mental retardation: Data from a population-based study. *Developmental Medicine and Child Neurology*, *42*(4), 266–270.

Taanila, A., Ebeling, H., Heikura, U., & Järvelin, M.-R. (2003). Behavioural problems of 8-year-old children with and without intellectual disability. *Journal of Pediatric Neurology*, *1*(1), 15–24.

Tassé, M. J., Aman, M. G., Hammer, D., & Rojahn, J. (1996). The Nisonger Child Behavior Rating Form: Age and Gender Effects and Norms. *Research in Developmental Disabilities*, *17*(1), 59–75.

Tassé, M. J., Morin, I. N., & Girouard, N. (2000). Traduction et validation canadienne-française du Nisonger Child Behavior Rating Form. *Canadian Psychology/Psychologie Canadienne*, *41*(2), 116–123. https://doi.org/10.1037/h0086861

Totsika, V., Hastings, R. P., Emerson, E., Lancaster, G. A., & Berridge, D. M. (2011). A population-based investigation of behavioural and emotional problems and maternal mental health: Associations with autism spectrum disorder and intellectual disability: ASD and intellectual disability. *Journal of Child Psychology and Psychiatry*, *52*(1), 91–99. https://doi.org/10.1111/j.1469-7610.2010.02295.x

Table 1

Expert Panel Members Degrees and Experience

Expert	**Terminal Degree**	**Years of Experience**	**Clinical/Research Experience**
1 (AW)	PhD	11	Specialized PhD in IDD psychology, licensed psychologist, focus of clinical/research in ID/ASD.
2 (AE)	PhD	15	Specialized PhD in IDD psychology, licensed psychologist, focus of clinical/research in ID/ASD.
3 (BMK)	PhD	23	Clinical psychology, licensed psychologist, clinical focus on neurodevelopmental disorders, Director of clinical internship program at a children's hospital.
4 (JW)	PhD	26	Specialized PhD in IDD psychology, licensed psychologist, focus of clinical practice in ID/ASD, Director of Clinical and Health Resources at ID/DD agency.
5 (MN)	PhD	10	Specialized PhD in IDD psychology, licensed psychologist, focus of clinical and research in ID/ASD – including psychometrics and assessment of psychopathology.
6 (AR)	MD, MPH	30	Licensed psychiatrist, areas of specialization include ID, DD, ASD. Provides clinical care to individuals with ID/ASD and comorbid psychiatric disorders.

Table 2

Existing NCBRF Items Matched to Diagnostic Criteria

Subscale	Existing NCBRF Item
ADHD	Short attention span
	Difficulty concentrating*
	Easily distracted*
	Fails to finish things that he/she starts
	Fidgets, wiggles, or squirms
	Overactive, doesn't sit still
	Talks too much or too loud
	Restless, high energy level*
Anxiety Disorders	Too fearful or anxious
	Worrying
	Restless, high energy level*
	Difficulty concentrating*
	Irritable
	Nervous or tense
	Feels others are against him/her
	Self-conscious or easily embarrassed
	Says no one likes him/her*
	Overly anxious to please others

Depressive Disorders	Underactive, slow
	Apathetic or unmotivated
	Restless, high energy level*
	Feels worthless or inferior
	Says no one likes him/her*
	Difficulty concentrating*
	Easily Distracted*
	Threatens to harm self
	Physically harms or hurts self on purpose
Oppositional Defiant	Temper tantrums
and Conduct Disorders	Easily frustrated
	Explosive, easily angered*
	Violates rules
	Lying or cheating
	Doesn't feel guilty after misbehaving
	Physically attacks people*
	Cruelty or meanness to others
	Gets in physical fights
	Steals
	Knowingly destroys property
Self-Injury and	Hits or slaps own head, neck, hands, or other body parts

Aggression	Harms self by scratching skin or pulling hair
	Gouges self, puts things in ears, nose, etc. or eats inedible things
	Repeatedly bites self hard enough to leave tooth marks or break skin
	Explosive, easily angered*
	Physically attacks people*

*_Note:_ these items were originally hypothesized to cross load on multiple subscales

Table 3

Delphi Method Results, Question Type 1

Question Type 1: Choose one: Item 1, Item 2, keep both times, or discard both items and write in a new item

Original NCBRF Items	**Round 1 Results**	**Round 2 Results: Cumulative item ranking**	**Round 3 Results: Agreement with top ranking item**
"Crying, tearful episodes" and "Unhappy or sad"	60% chose both items 40% chose "Unhappy or sad"	100% consensus to use both items	No feedback needed
"Isolates self from others" and "Withdrawn, uninvolved with others"	80% chose to write a new item 20% chose both items	1. New item: "Loss of interest in previously enjoyed activities" 2. New item: "Lack of Interest in what used to see fun" 3. New item: "Does not want to do things previously enjoyed" 4. Use both original items 5. New item: "Disinterested in typical (social) activities"	100% agree
"Shy around others; bashful" and "Shy or timid behavior"	60% chose to write a new item 20% chose "Shy or timid behavior"	1. New item: "Often feels embarrassed around others" 2. New item: "Shy around others"	83% agree

	20% chose "Shy around others; bashful"	3. New item: "Avoids social situations" 4. "Shy or timid behaviors" 5. "Shy around others; bashful"	
"Argues with parents, teachers, or other adults" and "Talks back to teachers, parents, or other adults"	60% chose "Argues with parents, teachers, or other adults" 20% chose "Talks back to teacher, parents, or other adults" 20% chose both items	1. "Argues with parents, teachers, or other adults" 2. Use both items 3. "Talks back to teacher, parents, or other adults"	100% agree
"Defiant, challenges adult authority" and "Disobedient"	60% chose both items 40% chose "Defiant, challenges adult authority"	1. Use both items 2. "Defiant, challenges adult authority" 3. Disobedient	100% agree
"Argues with other children or peers" and "Threatens people"	60% chose “Threatens people”* 40% chose both items*	1. "Threatens people" 2. New Item: "Bullies others" 3. Use both items 4. New Item: "Tries to dominate peers" 5. "Argues with other children or peers"	50% agree

**note: On this question, two panel members indicated two responses, instead of choosing one. This resulted in one panel members indicating “threatens people,” plus a new item they wrote, and a second panel member indicating both items plus a new item they wrote*

Table 4

Delphi Method Results, Question Type 2

Question Type 2: Presentation of New or Edited Item

Author-Written NCBRF-2 Item	**Round 1 Results:**	**Round 2 Results: Cumulative item ranking**	**Round 3 Results: Agreement with top ranking item**
"Unexpected change in appetite or weight"	60% chose item as is 40% revised the item	1. New Item: "Unexplained change in appetite or weight" 2. "Unexpected change in appetite or weight"	83% agree
"Sleep problems"	60% chose item as is 40% revised the item	1. "Sleep problems" 2. New Item: "Sleeps more or less than expected" 3. New Item: "Increase or decrease in sleep"	67% agree
"Disproportional fear to objects or situations (i.e., bees, heights)"	80% revised the item 20% chose item as is	1. New Item: "Excessive fear to objects or situations (i.e., bees, heights)" 2. New Item: "Afraid of many things" 3. "Disproportional fear to objects or situations (i.e., bees, heights)"	67% agree
"Gets tired easily"	100% chose item as is	No response needed	
"Sleep problems"	80% chose item as is 20% revised the item	1. "Sleep problems" 2. New Item: "Trouble sleeping"	67% agree
"Clings to adults"	60% chose item as is 40% revised the item	1. New Item: "Clings to adults in social situations" 2. "Clings to adults"	100% agree

"Refuses to talk in social situations"	80% chose item as is 20% revised the item	1. New Item: "Refuses to interact in social situations" 2. "Refuses to talk in social situations"	83% agree
"Shifts rapidly from topic to topic"	60% chose item as is 40% revised the item	1. New Item: "Shifts rapidly between topics or tasks" 2. New Item: "Shifts rapidly from activity to activity" 3. "Shifts rapidly form topic to topic"	83% agree
"Misplaces or loses things"	100% chose item as is	No response needed	No response needed
"Forgetful in routine activities"	80% chose item as is 20% revised the item	1. New Item: "Forgetful during routine activities" 2. "Forgetful in routine activities"	100% agree
"Has trouble waiting his/her turn (i.e., in conversation, standing in line)"	100% chose item as is	No response needed	No response needed
"Cruel or mean to animals"	100% chose item as is	No response needed	No response needed
"Runs away from home or school"	100% chose item as is	No response needed	No response needed
"Hits head on walls or doors"	80% chose item as is 20% revised item	1. New item: "Hits head on objects, walls, or doors" 2. "Hits head on walls or doors"	83% agree

Table 5

Delphi Method Results, Question Type 3

Question Type 3: Alignment with Subscale and Subsequent Revision

Original NCBRF Item	**Round 1 Results: Item Placement**	**Round 2 Results: Item Quality**	**Round 3 Results**
"Overly sensitive; feelings easily hurt"	80% chose to place item in Anxiety Subscale 20% chose to place item in Anxiety and Depression subscales	80% chose item as is 20% revised item	83% agree on original item "Overly sensitive; feelings easily hurt"
"Sudden changes in mood"	80% chose to discard item 20% chose to place item in Anxiety, Depression, and ODD subscales	No response needed; Item discarded	No response needed
"Sulks, is silent and moody"	60% chose to place item in Depression subscale 20% chose to place item in Depression and ODD subscales 20% chose to place item in Depression, Anxiety, and ODD subscales	80% chose item as is 20% revised item	New and original items tied for top ranking (i.e. “Sulks, is silent and moody,” and “Is silent or moody, mopes”)

Table 6

Delphi Method Results, Question Type 4

Question Type 4: Open-ended question to write in new items

Item Request	**Round 1 Results:**	**Round 2 Results:**	**Round 3 Results**
Request for additional "self-injury" items, if needed	60% did not write a new item 40% proposed a new item	New Item: "Doesn't stop hurting self when injured" 100% chose item as is	No response needed
		New Item: "Has injured self requiring medical attention" 80% chose item as is 20% chose to discard item	80% agree to use item
Request for additional "aggressive behavior" items, if needed	60% proposed a new item 40% did not write a new item	New Item: "Persistent use of epithets against peers or adults (e.g., disrespectful name calling and curse words)" 60% chose to discard item 40% revised item: Revised item: "Often curses at peers or calls them disrespectful names" Revised item: "Persistent in verbal abuse against peers or adults (e.g., disrespectful name calling and curse words)"	67% agree to discard item
		New Item: "Aggressive towards others (e.g., bites, scratches, slaps)"	No response needed

		100% chose item as is	

Table 7

NCBRF-2 Item Type Breakdown

Item Type	n	% of total n
Original and Unchanged NCBRF Items	43	67%
New Items Written by Authors	6	9%
New Items Written by Panel	15	23%
Total Items to be Included in Data Collection	64	

Table 8

Final Versions of New or Edited Item Stems Reviewed by Expert Panel

Subscale	Item after Delphi Method
ADHD	Shifts rapidly between topics or tasks
	Misplaces or loses things
	Forgetful during routine activities
	Has trouble waiting his/her turn (i.e., in conversations, standing in line)
Anxiety Disorders	Excessive fear to objects or situations (i.e., bees, heights)
	Gets tired easily
	Sleep problems*
	Often feels embarrassed around others
	Clings to adults in social situations
	Refuses to interact in social situations
Depressive Disorders	Loss of interest in previously enjoyed activities
	Unexplained change in appetite or weight
	Sleep problems*
	Is silent or moody, mopes
ODD/CD	Bullies others
	Cruel or mean to animals
	Runs away from home or school
Self-Injury and Aggression	Hits head on objects, walls, or doors
	Doesn't stop hurting self when injured

	Has injured self, requiring medical attention
	Aggressive towards others (e.g., bites, scratches, slaps)

Note: Items marked with (*) appear in more than one subscale