LIKERT-RATINGS VERSION OF THE DIPA

The original DIPA instrument rated the presence of symptoms dichotomously as either present or absent. This was in keeping with the Diagnostic and Statistical Manual strategy of endorsing symptoms as either present or absent. The psychometric study published in 2010 used this version (Scheeringa and Haslett, 2010). All versions of the DIPA prior to 2015 used these dichotomous ratings.

Since the original study, the DIPA has found traction as a useful measure at numerous research sites. Parts of or all of the DIPA have been used in over 20 peer-reviewed papers (list available from the author by request), and is being used at many clinical sites around the world.

In 2015, it was decided to change from dichotomous ratings to Likert-style ratings for every symptom. The original dichotomous responses of the DIPA limited its value to measure severity and to capture change over time. For example, a child could have five symptoms of depression and meet the full criteria for the diagnosis of major depressive disorder, and then receive treatment and show substantial reductions in all five symptoms. A post-treatment reassessment could theoretically show that all five symptoms were still present and the child still met full criteria for the diagnosis despite major improvements in severity and the family feeling satisfied with the treatment outcome. To expand the capacities of the DIPA beyond diagnoses to a more sensitive metric of change, the DIPA was revised to replace binary yes/no responses with 0-4 Likert-style ratings. This broader range of severity rating may have greater utility to determine prognosis, treatment planning, show change with treatment, and predict treatment response.

The Likert version has undergone psychometric testing to examine concurrent criterion validity and test-retest reliability. The results are in a manuscript under preparation.

MAJOR CHANGES FROM PRE-2013 VERSIONS OF THE DIPA

With the publication of the DSM-5 in May 2013, the DIPA has been thoroughly updated to include all relevant DSM-5 changes.

No changes were needed to MDD, ADHD, ODD, CD, anxiety, or sleep disorders.

PTSD

P10 Medical trauma
- “Life threatening” was added to definition.
- “…hospitalization, emergency room visit, and/or invasive medical procedures” was shortened to “…hospital visit or medical procedure” for brevity.

P11 Kidnapped
- This event was deleted because it was very rare in our experience. Replaced with “Learned that one of the events above happened to a caregiver” to match the DSM-5.

P14-P27 Acute reactions at the time of the events.
- Deleted to match the DSM-5.

Avoidance symptoms
- Avoidance of internal reminders swapped places with avoidance of external reminders to match the order in the DSM-5.
Avoidance of External Reminders
- “People” was moved from “avoidance of people, places, or things” to “avoidance of people, conversations, or interpersonal situations” to match the DSM-5.

Avoidance of Internal Reminders
- “avoidance of thoughts, feelings, conversations” was changed to “avoidance of people, conversations, or interpersonal situations” to match the DSM-5.

Inability to Recall Important Aspect of the Event
- Removed to match the DSM-5.

Increased Negative Emotional State
- Added to match the DSM-5.

Social Withdrawal and Restricted Affect switched order of places to match the order in the DSM-5.

Sense of Foreshortened Future
- Removed to match DSM-5.

Irritability Outbursts
- Aggression added to match DSM-5

Order of increased arousal symptoms moved around to match the order in the DSM-5.

Aggression
- Removed because it was subsumed under irritability symptom in DSM-5.

Disruptive mood dysregulation disorder
- This new disorder added in the depression section of the DSM-5.

Bipolar
B1. Excitement/energy
- The descriptor “felt like this” was changed to “had moods like this” to match the “mood” language in the DSM and to match with the new mood language in the grandiosity item.

Goal-directed activity or energy needed for criterion A in DSM-5. DSM-IV had required only the mood change for criterion A. DSM-5 required both mood and goal-directed activity. This did not change the DIPA itself, but required changing the algorithm on the DIPA Tally Sheet.

Reactive attachment disorder
Inhibited and Disinhibited subtypes from DSM-IV no longer used. Inhibited subtype simply became RAD. Disinhibited subtype became disinhibited social engagement disorder (DSED). The previous DIPA version had 5 items covering the old DSM-IV RAD, plus 3 items that covered
additional experimental symptoms. The newest DIPA version is split into 2 disorders – RAD (5 symptoms) and DSED (4 symptoms) – to match DSM-5.

An additional symptom was created by splitting “Overly familiar behavior, reduced reticence around unfamiliar adults” into 2 symptoms: “reduced reticence around unfamiliar adults” and “overly familiar verbal or physical behavior.” See below.

**Functional Impairment**
Functional impairment had never been required for the attachment disorders, so these questions were deleted.

**Disinhibited social engagement disorder**
This is a new disorder in DSM-5, created from the old disinhibited RAD subtype.

R6. Changed wording of “Does s/he act too familiar with stranger adults?” to “Does s/he approach unfamiliar adults too easily?” to make it more distinct from the new item on overly familiar behavior.

R7. Overly familiar verbal or physical behavior
This is a new item in DSM-5. It appears implicit that this is distinct from the symptom of reduced reticence around unfamiliar adults symptom (R6) by R6 being with unfamiliar adults and R7 including familiar adults.

**Anxiety**
Changed social phobia to social anxiety disorder to reflect the DSM-5 change in preference.

Clarification: In anxiety disorders, SP17 asks about onset and covers both specific phobia and social phobia. This was made clearer in the interview.

**Miscellaneous**
In the prior PTSD module, the directions for five symptoms stated “Probe with examples…same strategy as in P35.” Reference to P35 was correct in the 2011 version but was not correct in the 2013 version. These statements have been corrected to refer to item P15 or P21.

**TRANSLATIONS**
Translations of the original dichotomous DIPA have been completed in Arabic, Chinese, Danish, Dutch, French, German, and Spanish versions. These may be accessed at Dr. Scheeringa’s website [https://medicine.tulane.edu/departments/psychiatry/dr-scheeringas-lab/manuals-measures-trainings](https://medicine.tulane.edu/departments/psychiatry/dr-scheeringas-lab/manuals-measures-trainings)
DISORDERS COVERED

Posttraumatic stress disorder (PTSD)
Major depressive disorder (MDD)
Disruptive mood dysregulation disorder
Bipolar I disorder
Attention-deficit/hyperactivity disorder (ADHD)
Oppositional defiant disorder (ODD)
Conduct disorder (CD)
Separation anxiety disorder (SAD)
Specific phobia
Social phobia
Generalized anxiety disorder (GAD)
Obsessive compulsive disorder (OCD)
Reactive attachment disorder (RAD)
Disinhibited social engagement disorder
Sleep onset disorder
Night waking disorder

BASICS

The DIPA is intended as an interview for caregivers of children who are 6 years and younger down to below 1 year of age. It is not clear where the lower age limit is for the capacity to express each symptom because there are still limited data. It is acknowledged that many of the symptoms are not possible for infants but many scripts were worded so that they could be applied to younger children if desired and were not based on an a priori assumption that these symptoms could not be detected in younger children in the absence of data. Also, the DIPA ought to be easily extended to children older than 6 years if there were instances of desiring prospective continuity with the same measure for 7-12 year-old children.

Each disorder is in a self-contained module. All of the symptoms needed to make a DSM-5 diagnosis are in a disorder module and are presented in the same order that they are listed in the DSM-5 for ease of reference. Each module also contains a section for functional impairment, consistent with the DSM-5.

PHILOSOPHY

This instrument is meant for either lay or clinician interviewers. The format contains features of the two historical formats that have been called “structured” or “semi-structured.” Structured is the term that has been applied to interviews written for lay interviewers to read scripts in a fairly rigid fashion. This structure was designed for non-clinician interviewers employed in large-scale research surveys. Semi-structured has been applied to interviews written for interviewers to have more allowance to follow up initial probes with “free-lance” questions. This design was thought better-suited to clinicians who could use their training to track down relevant information.

The DIPA contains the structured characteristic that the interviewer is expected to read the first scripted probe for each item exactly as they are written. But overall, the DIPA is semi-structured because the interviewer is expected to follow-up every scripted probe with questions to obtain examples, and clarify examples until the interviewer is satisfied.
In actual practice, the distinction between structured and semi-structured is blurry because implementation of instruments depend more on the training and ongoing supervision of the interviewers than on how instruments were written. If lay interviewers giving structured interviews never understood some of the symptoms, drifted in their understanding of symptoms, and/or didn’t stick to the scripts, topics on which there are few data, then validity and standardization across subjects could suffer. Likewise, if clinicians using semi-structured interviews misunderstood items, used idiosyncratic phrasing, or collected incomplete data, then the same problems could occur. These important points are described in more detail below:

(1) The initial probe questions are detailed and ought to be read verbatim.

The initial probes questions (scripts) were made structured and detailed. These scripts are educational for both the interviewer and the respondent. The author’s experience over the years is that no matter how experienced the interviewers, there inevitably occur drift and misunderstandings. Drift and misunderstandings occur particularly if the scripts are left too vaguely worded. Therefore, the DIPA was created with relatively more detailed scripts than most diagnostic interviews in order to prevent drift and misunderstandings.

(2) To determine when an item passes the threshold of being a symptom, “the parent is the first judge, the interviewer is the final judge.”

One of the most important issues of validity is determining when an item passes the threshold of being a symptom. The DIPA philosophy of the “the parent is the first judge, the interviewer is the final judge” means that there is an explicit reliance on parents’ ratings but there must always be follow-up probing. All diagnostic interviews rely on respondents being the first judge, and sometimes the only judge if the interview is highly structured. In this sense, all interviews have relied on the respondents accessing a frame of reference in their minds to determine whether they consider the item in question abnormal or not. All clinicians and researchers are therefore somewhat dependent on the premise that all respondents have a common frame of reference of what is normal and abnormal. A unique feature of the DIPA is that it makes this common frame of reference explicit in the dialogue between the interviewers and the parents. Nearly all the DIPA scripts include the phrase that the item needs to be evident “more than the average child his/her age.” Other strategies in the DIPA are to ask explicitly if the parents consider these as “problems” or “excessive.”

These strategies both help to continually educate parents about what information is sought and shortens the interview by sidestepping conversations about normal emotions and behaviors. Another advantage is that this sidesteps having to write scoring rules for when each behavior crosses the threshold to be a symptom that would constrict the raters (interviewers) inappropriately in the absence of data in this age group. A disadvantage is that the assumption of a common frame of reference has limitations in the purest sense because of individual differences between caregivers. However, since interviewers must ask for an example for a behavior (“the interviewer is the final judge”) before it can be endorsed by the interviewers, this diminishes the risk that normative behaviors will be endorsed as problems.

SOME POINTS ABOUT PHRASING

The phrasing of the scripted questions is purposefully made more informal and conversational. This is meant to break up the monotony of repetitive phrasing and to make the
interview more tolerable for both respondents and interviewers. Some scripts state that certain phrases will not be repeated over and over, but interviewers would like respondents to remember the phrase for all questions. This applies to the “more than the average child his/her age” rule and the “in the last 4 weeks” rule. This is meant to reduce the stiltedness of the interview. Sometimes questions are turned into statements for variety and it is implicitly understood that it is still a question. Informal words are used that are more commonly used in normal conversation, such as “a bunch of things” instead of “multiple symptoms”; and instead of starting every question formally with “Does s/he do this item. . .”, it is assumed that the questioning pattern is implicit and the question starts only with “And how about this item. . .”. This interview also makes it explicit to respondents that some parts of the interview are redundant, and we openly apologize for it. This makes for good “customer service” so that the respondents feel like the interviewers understand the “inconvenience.”

There are three options in which the index child is referenced throughout the interview: (1) “your child”, (2) “s/he” (or “him/her”), and (3) “X”. “X” is to be replaced by interviewers with the child’s name. These three options are interspersed randomly for variety.

ACCEPTABLE INTERVIEW TECHNIQUES

All interview techniques are acceptable as long as they meet the main goals of the interview: (1) complete coverage of all items, (2) examples are obtained so that the interview can be the final judge as to whether to endorse an item, (3) the contextual frame is explicit that the behaviors are more than the average child his/her age, and (4) the interview is educational for the respondent when needed to help the respondent understand what is being asked.

Different techniques are needed for different types of respondents. The most common types of respondents tend to fall into three categories:

(1) Easy. Respondents quickly grasp the spirit of the questions, concisely answer the questions, and are not tangential. The interviewers can more or less read the scripts from the DIPA in sequential order and follow up linearly with probes.

(2) Moderate. Respondents don’t grasp easily the spirit of the questions. The interviewers need to explain items in more detail beyond what is provided in the scripts. Reading the scripts with little else would not extract the most accurate data.

(3) Difficult. Respondents are talkative and/or tangential. The interviewers may skip around, follow the respondents, and have to be more flexible about engaging in chitchat to maintain the cordiality of the interactions. Reading the scripts sequentially would clash.

SCORING

The new Likert version of the DIPA allows both categorical (i.e., diagnoses) and continuous (severity) scoring. **Categorical diagnoses scoring**

Symptoms are rated on 0-4 Likert scales. It is recommended that the presence of symptoms be categorically counted as present for any score of 1, 2, 3, or 4. It is acknowledged that others may argue that the presence of a symptom should be counted only for scores of 2, 3, or 4. However, based on experience with hundreds of caregivers of very young children, it has
been apparent that caregivers tend to underestimate the presence of symptoms due to the limited verbal capacities of young children. It seems wise to err on the side of counting something as a symptom even if it is rated as “1”.

There is a tally sheet at the beginning of the DIPA that includes the diagnostic scoring algorithms for all 16 disorders. This mirrors the DSM-5 diagnostic algorithms. This tally sheet can be used to determine if children meet diagnostic criteria for categorical disorders.

**Continuous (severity) scoring**

A continuous symptom severity score can be calculated for most disorders by summing the Likert ratings for each symptom. The functional impairment ratings should not be counted in the symptom severity scores.

Presently, there are no cutoffs for symptom severity scores that indicate levels of no concern, moderate concern, or probable diagnosis.

**Rounding for Frequency and Duration**

Caregivers often give ranges instead of precise numbers. **Take the average and round up** (if needed). For example, if a caregiver said 10 to 20 times for frequency, 15 would be entered. If a caregiver said 10 to 15 times, the average would be 12.5 and rounded up to 13.

**Frequency**

Frequency is usually coded as the number of instances over the **last four weeks**. This is a somewhat arbitrary period and can be changed to other periods if needed for other study designs. The two main reasons for choosing four weeks is that the majority of disorders in the DSM-5 require a four-week duration and this is a manageable period of time for caregivers to recall accurately.

For example, if the caregiver said “once a day, every day”, enter 28.

If a caregiver said three to four times a day, every day, then the lower bound would be three times 28, or 84. The upper bound would be four times 28, or 112. The average of 84 and 112 is 98.

Suppose a caregiver said five to six days a week, and “several times” on those days. First, make the caregiver clarify “several times.” Suppose the clarification leads to three times. The lower bound would be five days per week multiplied by four weeks, which equals 20; then multiply 20 by three times per day, which equals 60. The upper bound would be six times a week multiplied by four weeks, which equals 24; then multiply 24 by three times a day, which equals 72. The average of lower bound of 60 and the upper bound of 72 is 66.

An exception is the symptoms for CD which must be tracked over six months, not four weeks.

**Duration**

Durations are scored in minutes unless stated otherwise. If the duration is “all day”, this is counted as only the waking hours, or 16 hours (960 minutes). Two days would be 960 multiplied by 2 days, or 1,920 minutes. Two and one-half days would be 2,400 minutes, etc.

If the type of item has a duration that has no natural end to it, and only really ends because the parent did something to end it, or the child simply wore out and fell asleep, the duration scoring would be problematic.
Example: Refusal to sleep alone in separation anxiety disorder for most kids does not have an end. Children are continually afraid until the parent joins them or they fall asleep. It doesn’t make sense to code how long it takes the parent to join them or for them to fall asleep. Duration is not collected for this item.

Onset

Onset dates are recorded as mm/dd/yy.

Oftentimes, caregivers protest that they cannot remember the date something started. However, with patience and successive approximations, nearly every caregiver can remember dates with modest accuracy. Some suggestions for jogging caregivers’ memories include:

“Was it summer or winter?”
If the answer is winter, then ask, “Before Christmas or after Christmas?”
If the answer is summer, then ask, “During school or was school out?” (Even if the index child was not in school, there are often siblings in the home).

Major holidays are also useful landmarks:
“Was it before or after Easter?” (or July 4th or Thanksgiving).

Birthdays are also useful landmarks:
“Was it before or after his fourth birthday?”

If the caregiver says “early” in a month, code it as the 7th day. For example, for “early March” of 2006, code 03/07/06.
If the caregiver says “middle” of the month, code it as the 15th day.
If the caregiver says “later” in the month, code it as the 28th day.

If the item is not present and therefore there is no onset date, leave it blank or enter a “fake” date of your choice (e.g., 99/99/99).

FUNCTIONAL IMPAIRMENT

Consistent with the DSM-5 conceptualization of disorders, functional impairment must be assessed. This is assessed at the end of every module for that particular disorder. This is assessed with five questions about role functioning (with parents, with siblings, with peers, at school/day care, and in public) and one question about children experiencing distress from the symptoms (it is arguable whether child distress counts as functional impairment but these are discussed together here because they are lumped into a single item in each disorder in the DSM-5).

Linkage of Symptoms to Functional Impairment
For an impairment to be endorsed it must be clear that it was a change in functioning that was caused by the presence of symptoms. However, there are 2 exceptions to this rule:
(1) Lifelong symptoms

If trauma(s) has occurred since infancy and a symptom and its related functional impairment have always been present, then an impairment may be endorsed even though there has been no detectable temporal sequence of a change in functioning.
(2) Circumstance not encountered

If a particular circumstance (such as school) has not been encountered during the preceding four weeks, but it is clear that impairment was present in that circumstance the last time the child encountered it, then an impairment should be coded on the basis of the last time the child encountered it, assuming the caregiver agrees that this would accurately reflect the current impairment.

The DIPA has three unique questions of related significance in these sections: accommodations, perceived as a problem, and perceived need for treatment.

Accommodation

For four of the role functioning questions (with parents, with siblings, with peers, and in public), an extra question is added about whether parents make accommodations to keep the children out of those contexts specifically because they know it will lead to problems if they don’t. The purpose of this is to capture if children are not showing impairment because their caregivers have made adjustments. Therefore, accommodation questions are asked even if caregivers state that there is no impairment.

If respondents answer “no” to an impairment question, but “yes” to accommodation, we would count that as impairment.

Accommodation may be skipped if parents endorsed impairment. However, this is often an informative question to ask anyway for both research and clinical data-gathering.

Perceived as a Problem

Each impairment section contains one question about whether caregivers perceive the endorsed symptoms as problems. It is not uncommon with preschool children that parents will endorse much symptomatology and even impairment, but not perceive these things as problems, perhaps because young children are small, physically manageable, and/or parents believe children will grow out of it. This question provides a means to track this issue with data.

Perceived Need for Treatment

A second-tier issue of whether issues are perceived as problems is whether caregivers believe the problems need treatment. Caregivers might perceive that children have real problems, but still don’t believe that they need treatment. This question provides a means to track this issue with data.

DISORDER-SPECIFIC ISSUES

Bipolar I Disorder
Symptoms can be reviewed as lifetime instead of during the last 4 weeks time frame because bipolar is a uniquely episodic disorder.
A Manic Episode in DSM-5 is defined as 1 week in duration for the symptom of elevated, expansive, or irritable mood. This criterion, plus other aspects of bipolar disorder, has been controversial during the past decade. This interview is not meant as an endorsement of any criterion but rather is meant to be a data gathering tool to facilitate empirical solutions to the debate.

PTSD

Traumatic Life Events (page 1)
For a disaster, it would be useful to track children who directly experienced the event from those who only experienced the aftermath. For example, the convention used with the DIPA following Hurricane Katrina was to endorse item #4 “natural disasters” if children directly experienced the disaster (stayed through the storm). If children evacuated before the storm, and then returned later and saw their devastated homes, then endorse item #12 “other.”

Intrusive Recollections
The sub-item “Affect When Talking About It” is simply for data gathering about whether children are distressed during intrusive recollections (per the old DSM-IV requirement) or not (as has been empirically observed in research). Whether intrusive recollections are counted as a symptom contingent on distress or not is up to each researcher or clinician. The research of Scheeringa et al. that first documented this does not require distress for the symptom (Scheeringa, Peebles, Cook, & Zeanah, 2001; Scheeringa, Zeanah, Drell, & Larrieu, 1995; Scheeringa, Zeanah, Myers, & Putnam, 2003).

Endorse avoidance and/or distress at reminders even if there has been no exposure to reminders in the past four weeks, but the caregiver believes the symptom would still be manifest if there were to be an exposure today.

Children may have avoidance and/or distress at reminders but show no episodes in the last four weeks only because their caregivers have not exposed them to reminders. In these cases, interviewers must ask, “If your child were to be in such a situation (anticipation of exposure for avoidance or actual exposure for distress at reminders), do you think s/he would show the behavior (avoidance or distress)?” If caregivers believe that children would still show the behavior today, then the appropriate item must be endorsed.

Fears that are part of distress at reminders, avoidance, and other new fears.
Fears that are discussed in the PTSD section, when viewed in isolation, appear identical to the fears of specific phobia. Fears ought not to be coded in more than one disorder module. If the onset of the fear appears related to a traumatic event, then code it under PTSD. If not, code it under specific phobia.

Phobias
Frequency is not rated for phobias because these are driven by chance and circumstance of environmental exposures.

Both specific phobia and social phobia have an A criterion that is the “fear trait” aspect and a B criterion that is the “anxiety response to exposure” aspect. In young children, A often can only be inferred by the presence of B. Therefore, A is not really asked in this interview but B is. If B is present, this implicitly suggests that A is present, although the limitations of such inferences are acknowledged.

The C criterion that the fear must be recognized by children as excessive or unreasonable is not asked due to developmental limitations.

PSYCHOMETRICS

The test-retest reliability and concurrent criterion validity properties of the original categorical DIPA were tested in a research study with 1-5 year old children for seven disorders - PTSD, MDD, ADHD, ODD, SAD, GAD, and OCD (Scheeringa & Haslett, 2010). The DIPA showed
adequate properties but there were too few children with GAD and OCD symptoms to fully test those two disorders. The details were published in:

Psychometric properties of the newer Likert version are in a manuscript under preparation.

FAQ

Doesn’t an interview like this create a sterile atmosphere and parents won’t open up about their children’s personal issues?

For most parents, no. Most respondents view the structure of the interview as a good quality. To make it more “customer service” friendly however, it is recommended to say explicitly at the beginning that the interviewer needs to follow this interview and it will take a long time. For respondents who find the interview uncomfortable, this is usually a reflection of larger issues that are not isolated to this interview. Also, one must rely on the interpersonal skills of the interviewers to insert informality and conversational free-lancing as needed on an individual basis.

Why is the script “and this was present sometime in the last 4 weeks?” written out for nearly every item? Why can’t you assume interviewers will remember this and save ink?

In past experience, interviewers don’t remember this. The interview is sufficiently complicated that this seems to be forgotten fairly regularly and items get coded for lifetime if this reminder is not constantly written in the interview.

Why aren’t scripts written out for every item to remind the interviewer to ask about frequencies and durations?

There are blank lines with labels in the far right column for frequency and duration for every item where applicable. These appear to be sufficient reminders for interviewers.

If core, required criteria are not met, why aren’t there skip outs to the next disorder module?

This is relevant only to PTSD and major depression. Most other published diagnostic interviews have a skip out to the next module if, for example, an individual does not have sadness or loss of interest under major depression. The decision was made to capture information about all possible items for these disorders, not just to determine the categorical presence of a disorder.
LITERATURE CITED


