



Published in final edited form as:

Biol Psychiatry. 2016 June 15; 79(12): e91–e96. doi:10.1016/j.biopsych.2015.07.006.

Common Measures for National Institute of Mental Health Funded Research

Deanna M. Barch, Ian H. Gotlib, Robert M. Bilder, Daniel S. Pine, Jordan W. Smoller, C. Hendricks Brown, Wayne Huggins, Carol Hamilton, Adam Haim, and Gregory K. Farber

Departments of Psychology, Psychiatry, and Radiology (DMB), Washington University, St. Louis, Missouri; Department of Psychology (IHG), Stanford University, Stanford, California; Departments of Psychiatry and Biobehavioral Sciences and Psychology (RMB), University of California at Los Angeles, Los Angeles, California; National Institute of Mental Health (DSP), Intramural Research Program, Bethesda, Maryland; Psychiatric and Neurodevelopmental Genetics Unit (JWS), Center for Human Genetic Research, and Department of Psychiatry (JWS), Massachusetts General Hospital; and Stanley Center for Psychiatric Research (JWS), Broad Institute, Boston, Massachusetts; Department of Psychiatry (CHB), Northwestern University, Chicago, Illinois; RTI International (WH, CH), Research Triangle Park, Durham, North Carolina; and Office of Clinical Research (AH) and Office of Technology Development and Coordination (GKF), National Institute of Mental Health, Bethesda, Maryland.

One of the most encouraging, but also the most challenging, aspects of current research on psychopathology is the diversity of measures used to assess constructs across research studies and programs. Clearly, this diversity reflects the creativity and generativity of our field and the continual growth of our science. At the same time, however, this diversity also makes data harmonization across studies difficult, if not sometimes impossible. The National Human Genome Research Institute recognized this conundrum in the field of genetics and started an initiative referred to as consensus measures for Phenotypes and eXposures (PhenX) to identify and recommend a small number of measures for each of 21 broad research domains that could be used as common assessments to facilitate integration across genomewide association studies (1–4). These measures are made available to the scientific community, at no cost, in the PhenX Toolkit (<https://www.phenxtoolkit.org>). Subsequently, the PhenX consensus process was used to identify measures in support of substance abuse and addiction (SAA) research, adding depth to the toolkit in this area. This project was funded by the National Institute on Drug Abuse (NIDA) with the participation of the National Institute on Alcohol Abuse and Alcoholism. Perhaps due to a growing awareness of the need to share data across studies to increase statistical power and study impact, a number of other common data element programs have been underway, including the Patient-Reported Outcomes Measurement Information System (5), the National Institutes of Health (NIH) Toolbox (6), the Neurological Quality of Life (7), the National Institute of Neurological Disorders and Stroke Common Data Elements program (8,9), and the NIH

Address correspondence to Deanna M. Barch, Ph.D., Washington University, Departments of Psychology, Psychiatry and Radiology, Box 1125, One Brookings Drive, St. Louis, MO 63130; dbarch@wustl.edu.

Disclosures

Dr. Pine, Dr. Haim, and Dr. Farber report no biomedical financial interests or potential conflicts of interest.

Common Data Elements program (<http://www.nlm.nih.gov/cde/>). The program staff at the National Institute of Mental Health (NIMH), as well as its funded researchers, have also recognized the challenges posed by a lack of common measures across studies. The NIMH has taken note of this recent emphasis on larger scale studies to address core questions about the mechanisms of psychopathology and recent attempts at data harmonization across studies of psychopathology that address similar issues. Accordingly, the NIMH felt that it was time to identify brief, low-burden measures that NIMH-funded researchers could include in their studies to increase cross-study data compatibility. The goal of the current report is to briefly describe the genesis and development of the PhenX project, to outline the process that the Mental Health Research Panel used to select a set of common measures, to describe the measures themselves, and to outline the goals associated with including these measures in future studies.

PhenX

The goal of the original PhenX project, led by RTI International and funded by the National Human Genome Research Institute and the NIH Office of Behavioral and Social Sciences Research, was to identify and recommend well-established, low-burden, high-quality measures for use in large-scale genomic studies. The approach was to select up to 15 measures for each of 21 research domains and make them freely available to the research community via the PhenX Toolkit (<https://www.phenxtoolkit.org>). A 12-member steering committee decided the 21 domains to be addressed, including demographics, anthropometrics, environmental exposures, speech and hearing, psychiatric, and psychosocial, and set forth criteria for selecting measures (<https://www.phenxtoolkit.org/index.php?pageLink=browse>) (1). For each domain, a working group of six to nine experts was convened to identify measures using a rigorous consensus process that includes a series of conference calls, in-person meetings, and outreach to the research community (10). The measures selected by the original 21 working groups were completed in 2011 and resulted in 295 measures being released in the PhenX Toolkit.

PhenX Measures for Substance Abuse and Addiction Research

NIDA subsequently provided funding for a project to add depth to the toolkit in support of SAA research. The impetus for this project relates to the fact that there is very low commonality in the measures used across studies of abuse and addiction (11). This project, which added 45 measures to the PhenX Toolkit, included a core collection and six specialty collections of SAA measures. The SAA core collection consists of two tiers: tier 1 and tier 2. The Core: Tier 1 collection was designed to take respondents approximately 8 to 10 minutes to complete and includes basic demographic measures and assessments of substance use. Core: Tier 2 is expected to take approximately 27 to 46 minutes and includes expanded demographic information, substance use, psychiatric disorders screening, peer use of substances, and social network information.

NIDA's expectations for the inclusion of these core and specialty measures in studies funded by this agency are described in a published notice (<http://grants.nih.gov/grants/guide/notice-files/NOT-DA-12-008.html>):

“Core: Tier 1: The measures in this collection are deemed relevant and essential to all areas of addiction science. NIDA grantees/applicants conducting human-subject studies are strongly encouraged to incorporate, at a minimum, the Core-Tier 1 measures.

Core: Tier 2: The measures in this collection are deemed relevant to all areas of addiction science. Because Tier-2 measures are considered more burdensome and specialized than the Tier-1 measures, NIDA grantees/applicants conducting human-subject studies are strongly encouraged to incorporate them whenever possible and appropriate.

Specialty: The measures in this collection are deemed relevant and essential within specific areas of addiction science. NIDA grantees/applicants conducting human-subject studies in the specified areas of science are strongly encouraged to incorporate the Specialty measures.”

Since the SAA collections of measures were released in the PhenX Toolkit in the spring of 2012, both NIDA and NIAA have encouraged use of PhenX measures in numerous funding opportunity announcements.

NIMH Initiatives

An even broader emphasis in NIMH on cross-study harmonization began to emerge when NIMH encouraged eight different groups of researchers studying the prodrome of psychosis to combine their efforts into the North American Prodrome Longitudinal Study. The North American Prodrome Longitudinal Study was originally eight independently designed and funded projects (12) but suffered from the challenges associated with the relatively low conversion rates within any one relatively small sample, which could be bolstered by aggregating across samples. This data harmonization effort, while successful in many ways (13–18), illustrated the difficulties of combining data across studies using different measures to assess similar constructs. Thus, in 2013, the NIMH, in collaboration with RTI International, convened the Mental Health Research Panel (MHRP) to guide the selection of a core set of measures for mental health related research, similar to the Core: Tier 1 and Core: Tier 2 measures selected by the NIDA SAA scientific panel. Based on calls in the National Research Action Plan (http://www.whitehouse.gov/sites/default/files/uploads/nrap_for_eo_on_mental_health_august_2013.pdf), the NIMH has convened specialty working groups to identify and recommend measures focused on specific mental health related research questions (e.g., suicide, posttraumatic stress disorder). In addition, the NIMH has supported data integration approaches through the establishment of the National Database for Autism Research (19), a platform that is expanding to include data from Research Domain Criteria (RDoC) focused projects.

The MHRP recognized the tradeoffs associated with recommending that all NIMH-funded studies include a set of core measures. On the positive side, the inclusion of at least a small set of common measures will clearly facilitate cross-study replication and will allow the aggregation of data across studies into larger data sets that should permit statistically stronger tests of hypotheses or the testing of novel hypotheses that could not be addressed in any single study. On the negative side, any requirement that investigators include a specific

set of measures has to be balanced with the important need for researchers to choose and include measures of the specific constructs of interest to their particular studies. As such, some could view the emphasis on common data elements as hindering scientific innovation, though clearly both are important for forward progress. For these reasons, the MHRP felt that the recommended NIMH Core: Tier 1 and Core: Tier 2 measures needed to be very brief and low burden, ensuring that investigators would have ample time to include the additional measures they feel are important to addressing their specific aims. Consequently, we selected a set of measures designed to be primarily self-report and/or easy to administer and to take less than 15 minutes to complete (see Table 1 and the PhenX website, www.phenxtoolkit.org/, for a more detailed description of criteria and measure attributes). The MHRP also separately identified measures that would be appropriate for use in studies of adults/adolescents versus studies of children. Like NIDA, NIMH suggests that the Core: Tier 1 measures are relevant and essential to all areas of mental health research, and all NIMH grantees/applicants conducting human research will be strongly encouraged to include Core: Tier 1 measures in their studies. While the Core: Tier 2 measures are also relevant to all areas of mental health research, they are slightly more burdensome and specialized than the Core: Tier 1 measures. Thus, NIMH grantees/applicants conducting research with human subjects are strongly encouraged to incorporate Core: Tier 2 measures when feasible and appropriate.

NIMH Core: Tier 1

As shown in Table 1, several of the recommended NIMH Core: Tier 1 measures are focused on demographic and socioeconomic-related measures. We recognize that many, if not most, researchers already collect such measures in their studies. Importantly, however, the wording of the assessments and the ways in which they are coded differ significantly across studies. Thus, adopting a common framework would facilitate cross-study data aggregation and address questions related to the effects of factors such as age, race, and socioeconomic status (SES) on relevant outcomes. The goal of including child-reported parental education attainment (when the child is old enough to provide such reports) is to garner information on the role of developmental SES, given the growing body of data showing that developmental SES can have a major impact on both psychosocial and biological outcomes (20–23). In addition, the MHRP suggested that investigators include brief self-report measures of broad psychopathology and impairment for both child and adult studies. Some of these measures were not previously included in the PhenX, but we felt that their inclusion was critical for studies of mental health outcomes to provide a broad and basic assessment of these domains.

For a broad psychopathology assessment for children, we recommended including the Strengths and Difficulties Questionnaire, a brief, valid, and reliable measure that covers a broad range of behavior relevant to a number of different domains of symptoms and functioning (24–29). There are parent-report versions of this measure for 2- to 4-year-olds, 4- to 10-year-olds, and 11- to 17-year-olds, as well as a self-report version for children ages 11 to 17. For adults, we recommended including the DSM-5 Self-Rated Level 1 Cross-Cutting Symptom Measure. The MHRP had a spirited debate about including this measure, given the recent emphasis by NIMH on the RDoC initiative and the potential perception that recommending a DSM-5 oriented measure would be seen as counter to this initiative.

Nevertheless, most experts agree that DSM-5 and RDoC provide complementary approaches to assessment. As such, it is important to have a common, but brief, symptom-based measure in studies focused on mental health, even for studies that were generated primarily from an RDoC-based framework. We should also note, however, that the MHRP felt strongly that the question designed to assess suicidal ideation on this measure (question VI) was clearly not adequate for the assessment of the full range of suicide-related ideation; it felt that it may be better labeled self-harm. Those researchers specifically interested in a comprehensive assessment of suicide should see the forthcoming recommendations of the suicide working group.

For a measure of impairment in children, we recommended the use of the 13-item Columbia Impairment Scale (30), a brief validated measure of function in four areas: interpersonal relationships, psychopathology, school or work function, and leisure time. We recognize that an ideal measure would complete separate indicators of psychopathology from indicators of function but also realize that this is difficult, if not impossible, to accomplish in children. The Columbia Impairment Scale has both parent- and child-report versions. For adults, we recommended the 12-item self-report version of the World Health Organization Disability Assessment Schedule 2.0 (31–33).

NIMH Core: Tier 2

Given the broad range of research questions addressed in NIMH-funded studies, the MHRP found it challenging to identify additional measures that were relevant to all NIMH-funded, human-subjects research. However, three additional areas appear relevant to the vast majority of mental health related studies: assessments of substance use, life events, and perceived stress. Existing research clearly documents significant comorbidity of substance use difficulties in individuals who are experiencing mental health related difficulties. Thus, the MHRP recommended including the brief SAA Core: Tier 1 assessments of tobacco, alcohol, and substance use (11) in the NIMH Core: Tier 2, with separate measures for adults and adolescents. The PhenX Toolkit already includes self/parent report measures of life events for children [Adverse Life Events Scale (34,35)] and adults [Traumatic Life Events Questionnaire (36,37)], each of which had originally been selected by the social environments working group. We are aware that there is evidence that more intensive interview-based assessments of life events have incremental validity, but we felt that it was not feasible to recommend such measures in this context. For perceived stress, the PhenX Toolkit already includes the 10-item Perceived Stress Scale (38,39), selected by the psychosocial working group of experts who consulted broadly. The MHRP recognized that perceptions of stress can be influenced by current mood and other factors but felt that the predictive utility of assessing perceived stress was sufficient to mitigate against these concerns. For children, we could not identify a validated measure of self-reported perceived stress appropriate for pediatric populations or parent reports of their child's perception of stress. Thus, developing a measure that assesses such constructs appears to be an important task for future work.

Time Frames

One of the challenges in the use of the measures recommended for NIMH Core: Tier 1 and Core: Tier 2 is that the recommended time frames for assessment vary across measures. For example, whereas the World Health Organization Disability Assessment Schedule 2.0 asks about the last 30 days, the DSM-5 Level 1 Cross-Cutting measure inquires about the last 2 weeks and the Strengths and Difficulties Questionnaire asks about the last 6 months. Further, the appropriate time frame may vary for different types of studies (e.g., cross-sectional, longitudinal, treatment-focused). To preserve comparability across existing studies, the MHRP recommended that the default be to use the time frame recommended by the developers of the measures for all cross-sectional studies and for the baseline assessments for longitudinal and treatment studies, but the MHRP recognized that other time frames might be justified for projects in which the default time frame might be inappropriate or confuse participants. If these measures are used at subsequent time points in longitudinal or treatment studies, it is further recognized that researchers may need to modify the time frames in ways that are appropriate for their particular study design given the time intervals between assessments.

Opportunities and Challenges of Data Harmonization

As with most things, there are opportunities and challenges associated with a move toward even a small subset of common data elements. On the opportunities side, the use of common measures of broad psychopathology and impairment, as well as demographic factors, will support the aggregation of data sets of various types (e.g., structural and functional imaging, behavioral measures, etc.) that might then have greater power to detect, for example, interactions between demographic factors and psychopathology/impairment or between different aspects of psychopathology. Further, data aggregation will allow greater power in exploratory studies; the challenges of multiple comparison correction for many statistical tests can often preclude the generation of informative results that can guide future research. On the challenges side, sometimes a push toward common data elements can encourage investigators to use less than optimal measures [e.g., PhenX recommendation of digit span as a measure of working memory, which is not a state-of-the-art measure of all aspects of working memory (40)]. For these reasons, the MHRP focused on keeping the measures recommended for Core: Tier 1 and Core: Tier 2 very brief and focused on elements with the clearest clinical significance (e.g., psychopathology, impairment, substance use, demographics), leaving investigators to select what they believe are the state-of-the-art measures for the specific constructs of interest in their studies. We also recognize that there may be domains in which there are large and consistent effect sizes, and thus data aggregation may not be necessary to generate clinically meaningful and impactful data. At the same time, it should be recognized that greater consistency of measurements and increased ability to combine, compare, and/or contrast findings across studies may yield benefits beyond those that can be anticipated at the time of study.

Conclusions

The large-scale aggregation of data, beyond those feasibly collected by individual investigators, is a crucial engine for the discovery and validation of advances in mental health research. Perhaps the most striking proof of concept of this has come from genomic studies in which widespread data sharing, through the international Psychiatric Genomics Consortium, has transformed psychiatric genetics. Previous single investigator/group studies were limited by the challenges of multiple statistical tests with relatively small N s, and thus data aggregation has taken the field from having virtually no established risk loci in 2008 to nearly 200 confirmed loci today. This has included more than 100 common risk variants for schizophrenia alone (41). In fact, the availability of common data elements beyond simply diagnostic categories has become a rate-limiting step for further discovery.

We expect the PhenX Toolkit will be an important platform for enabling such discoveries in mental health, clinical neuroscience, and genetic research by facilitating data harmonization across studies. Certainly, there are numerous other opportunities to identify common measures that can be used across studies relevant to mental health, but relatively few of these measures would be relevant to all mental health related studies being funded by the NIMH. Instead, the MHRP felt that the next step was to examine such opportunities for common data elements at the level of subgroups of studies with shared goals and constructs, perhaps at the level of RDoC domains or constructs. As the first wave of RDoC-focused research comes to fruition, the validity of potential measures will become more apparent and the time may then be ripe for efforts focused on identifying common data elements for assessing RDoC constructs. An additional major opportunity created by increased harmonization of measurements across NIMH projects is the capacity to develop shared knowledge bases, which ultimately can be mined to address questions impossible on the scale of any single study.

Acknowledgments

Deanna Barch is funded by the National Institute of Mental Health (NIMH) (Grant Nos. R01MH66031, R01MH084840, R01MH090786, and R01MH098454-S), the National Institutes of Health Blueprint (Grant No. U54MH091657), and National Alliance for Research on Schizophrenia and Depression. Deanna Barch consults for Pfizer, Amgen, and Roche and has a contract to analyze imaging data for Pfizer. Dr. Gotlib is funded by NIMH (Grant Nos. MH074849 and MH101495). Dr. Bilder is funded by NIMH (Grant No. MH101478). Dr. Bilder has received honoraria or consulting fees from EnVivo, Forum Pharmaceuticals, Lumosity, Snapchat, Takeda-Lundbeck, and ThinkNow, Inc. Dr. Smoller is funded by NIMH (Grant No. MH094614) and is the Tepper Family Massachusetts General Hospital Research Scholar. Dr. Brown is funded by NIMH (Grant No. MH040859). Dr. Huggins is funded by the National Human Genome Research Institute (Grant No. U41 HG007050). Dr. Hamilton is funded by the National Human Genome Research Institute (Grant No. U41 HG007050). Dr. Smoller is an unpaid member of the Scientific Advisory Board of PsyBrain, Inc.

References

1. Hamilton CM, Strader LC, Pratt JG, Maiese D, Hendershot T, Kwok RK, et al. The PhenX Toolkit: Get the most from your measures. *Am J Epidemiol*. 2011; 174:253–260. [PubMed: 21749974]
2. McCarty CA, Huggins W, Aiello AE, Bilder RM, Hariri A, Jernigan TL, et al. PhenX RISING: Real world implementation and sharing of PhenX measures. *BMC Med Genomics*. 2014; 7:16. [PubMed: 24650325]

3. Pan H, Tryka KA, Vreeman DJ, Huggins W, Phillips MJ, Mehta JP, et al. Using PhenX measures to identify opportunities for cross-study analysis. *Hum Mutat.* 2012; 33:849–857. [PubMed: 22415805]
4. Stover PJ, Harlan WR, Hammond JA, Hendershot T, Hamilton CM. PhenX: A toolkit for interdisciplinary genetics research. *Curr Opin Lipidol.* 2010; 21:136–140. [PubMed: 20154612]
5. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap cooperative group during its first two years. *Med Care.* 2007; 45:S3–S11. [PubMed: 17443116]
6. Gershon RC, Cella D, Fox NA, Havlik RJ, Hendrie HC, Wagster MV. Assessment of neurological and behavioural function: The NIH Toolbox. *Lancet Neurol.* 2010; 9:138–139. [PubMed: 20129161]
7. Perez L, Huang J, Jansky L, Nowinski C, Victorson D, Peterman A, Cella D. Using focus groups to inform the Neuro-QOL measurement tool: Exploring patient-centered, health-related quality of life concepts across neurological conditions. *J Neurosci Nurs.* 2007; 39:342–353. [PubMed: 18186419]
8. Saver JL, Warach S, Janis S, Odenkirchen J, Becker K, Benavente O, et al. Standardizing the structure of stroke clinical and epidemiologic research data: The National Institute of Neurological Disorders and Stroke (NINDS) Stroke Common Data Element (CDE) project. *Stroke.* 2012; 43:967–973. [PubMed: 22308239]
9. Loring DW, Lowenstein DH, Barbaro NM, Fureman BE, Odenkirchen J, Jacobs MP, et al. Common data elements in epilepsy research: Development and implementation of the NINDS epilepsy CDE project. *Epilepsia.* 2011; 52:1186–1191. [PubMed: 21426327]
10. Maiese, DR.; Hendershot, T.; Strader, LC.; Wagener, D.; Hammond, JA.; Huggins, W., et al. PhenX: Establishing a Consensus Process to Select Common Measures for Collaborative Research. Research Triangle Park, NC: RTI Press; 2013. RTI Press Publication NO MR-0027-1310
11. Conway KP, Vullo GC, Kennedy AP, Finger MS, Agrawal A, Bjork JM, et al. Data compatibility in the addiction sciences: An examination of measure commonality. *Drug Alcohol Depend.* 2014; 141:153–158. [PubMed: 24954640]
12. Addington J, Cadenhead KS, Cannon TD, Cornblatt B, McGlashan TH, Perkins DO, et al. North American Prodrome Longitudinal Study: A collaborative multisite approach to prodromal schizophrenia research. *Schizophr Bull.* 2007; 33:665–672. [PubMed: 17255119]
13. Addington J, Cornblatt BA, Cadenhead KS, Cannon TD, McGlashan TH, Perkins DO, et al. At clinical high risk for psychosis: Outcome for nonconverters. *Am J Psychiatry.* 2011; 168:800–805. [PubMed: 21498462]
14. Cadenhead KS, Addington J, Cannon T, Cornblatt B, McGlashan T, Perkins D, et al. Treatment history in the psychosis prodrome: Characteristics of the North American Prodrome Longitudinal Study Cohort. *Early Interv Psychiatry.* 2010; 4:220–226. [PubMed: 20712727]
15. Cannon TD, Cadenhead K, Cornblatt B, Woods SW, Addington J, Walker E, et al. Prediction of psychosis in youth at high clinical risk: A multisite longitudinal study in North America. *Arch Gen Psychiatry.* 2008; 65:28–37. [PubMed: 18180426]
16. Seidman LJ, Giuliano AJ, Meyer EC, Addington J, Cadenhead KS, Cannon TD, et al. Neuropsychology of the prodrome to psychosis in the NAPLS consortium: Relationship to family history and conversion to psychosis. *Arch Gen Psychiatry.* 2010; 67:578–588. [PubMed: 20530007]
17. Velthorst E, Nieman DH, Becker HE, van de Fliert R, Dingemans PM, Klaassen R, et al. Baseline differences in clinical symptomatology between ultra high risk subjects with and without a transition to psychosis. *Schizophr Res.* 2009; 109:60–65. [PubMed: 19272756]
18. Woods SW, Addington J, Cadenhead KS, Cannon TD, Cornblatt BA, Heinssen R, et al. Validity of the prodromal risk syndrome for first psychosis: Findings from the North American Prodrome Longitudinal Study. *Schizophr Bull.* 2009; 35:894–908. [PubMed: 19386578]
19. Hall D, Huerta MF, McAuliffe MJ, Farber GK. Sharing heterogeneous data: The national database for autism research. *Neuroinformatics.* 2012; 10:331–339. [PubMed: 22622767]

20. Luby J, Belden A, Botteron K, Marrus N, Harms MP, Babb C, et al. The effects of poverty on childhood brain development: The mediating effect of caregiving and stressful life events. *JAMA Pediatr.* 2013; 167:1135–1142. [PubMed: 24165922]
21. Luby JL, Barch DM, Belden A, Gaffrey MS, Tillman R, Babb C, et al. Maternal support in early childhood predicts larger hippocampal volumes at school age. *Proc Natl Acad Sci U S A.* 2012; 109:2854–2859. [PubMed: 22308421]
22. Butterworth P, Cherbuin N, Sachdev P, Anstey KJ. The association between financial hardship and amygdala and hippocampal volumes: Results from the PATH through life project. *Soc Cogn Affect Neurosci.* 2012; 7:548–556. [PubMed: 21551226]
23. Hanson JL, Chandra A, Wolfe BL, Pollak SD. Association between income and the hippocampus. *PLoS One.* 2011; 6:e18712. [PubMed: 21573231]
24. Bourdon KH, Goodman R, Rae DS, Simpson G, Koretz DS. The Strengths and Difficulties Questionnaire: U.S. normative data and psychometric properties. *J Am Acad Child Adolesc Psychiatry.* 2005; 44:557–564. [PubMed: 15908838]
25. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry.* 2001; 40:1337–1345. [PubMed: 11699809]
26. Goodman R, Meltzer H, Bailey V. The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *Eur Child Adolesc Psychiatry.* 1998; 7:125–130. [PubMed: 9826298]
27. Kovacs S, Sharp C. Criterion validity of the Strengths and Difficulties Questionnaire (SDQ) with inpatient adolescents. *Psychiatry Res.* 2014; 219:651–657. [PubMed: 25048754]
28. Muris P, Meesters C, Eijkelenboom A, Vincken M. The self-report version of the Strengths and Difficulties Questionnaire: Its psychometric properties in 8- to 13-year-old non-clinical children. *Br J Clin Psychol.* 2004; 43:437–448. [PubMed: 15530213]
29. Muris P, Meesters C, van den Berg F. The Strengths and Difficulties Questionnaire (SDQ)—further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *Eur Child Adolesc Psychiatry.* 2003; 12:1–8. [PubMed: 12601558]
30. Bird HR, Canino GJ, Davies M, Ramirez R, Chavez L, Duarte C, Shen S. The Brief Impairment Scale (BIS): A multidimensional scale of functional impairment for children and adolescents. *J Am Acad Child Adolesc Psychiatry.* 2005; 44:699–707. [PubMed: 15968239]
31. Gold LH. DSM-5 and the assessment of functioning: The World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0). *J Am Acad Psychiatry Law.* 2014; 42:173–181. [PubMed: 24986344]
32. Konecky B, Meyer EC, Marx BP, Kimbrel NA, Morissette SB. Using the WHODAS 2.0 to assess functional disability associated with DSM-5 mental disorders. *Am J Psychiatry.* 2014; 171:818–820. [PubMed: 25082488]
33. Ustun TB, Chatterji S, Kostanjsek N, Rehm J, Kennedy C, Epping-Jordan J, et al. Developing the World Health Organization Disability Assessment Schedule 2.0. *Bull World Health Organ.* 2010; 88:815–823. [PubMed: 21076562]
34. Grant KE, Compas BE, Thurm AE, McMahon SD, Gipson PY. Stressors and child and adolescent psychopathology: Measurement issues and prospective effects. *J Clin Child Adolesc Psychol.* 2004; 33:412–425. [PubMed: 15136206]
35. Tiet QQ, Bird HR, Hoven CW, Moore R, Wu P, Wicks J, et al. Relationship between specific adverse life events and psychiatric disorders. *J Abnorm Child Psychol.* 2001; 29:153–164. [PubMed: 11321630]
36. Kubany ES, Haynes SN, Leisen MB, Owens JA, Kaplan AS, Watson SB, Burns K. Development and preliminary validation of a brief broad-spectrum measure of trauma exposure: The Traumatic Life Events Questionnaire. *Psychol Assess.* 2000; 12:210–224. [PubMed: 10887767]
37. Peirce JM, Burke CK, Stoller KB, Neufeld KJ, Brooner RK. Assessing traumatic event exposure: Comparing the Traumatic Life Events Questionnaire to the Structured Clinical Interview for DSM-IV. *Psychol Assess.* 2009; 21:210–218. [PubMed: 19485675]
38. Cohen S, Janicki-Deverts D. Who's stressed? Distributions of psychological stress in the United States in probability samples from 1983, 2006 and 2009. *Appl Soc Psychol.* 2012; 42:1320–1334.

39. Cohen S. Contrasting the hassle scale and the perceived stress scale. *Am Psychol.* 1986; 41:716–719.
40. Gold JM, Hahn B, Zhang WW, Robinson BM, Kappenman ES, Beck VM, Luck SJ. Reduced capacity but spared precision and maintenance of working memory representations in schizophrenia. *Arch Gen Psychiatry.* 2010; 67:570–577. [PubMed: 20530006]
41. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014; 511:421–427. [PubMed: 25056061]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1**Mental Health Research Measures**

Measure	Description of Protocol
1. Current Age	The respondent is asked for his/her date of birth in MM/DD/YYYY format. If the respondent does not know his/her date of birth, there is an alternative question to determine his/her age.
2. Race	The respondent is asked to indicate one or more categories that describe what race he or she considers himself or herself to be. The question meets the Office of Management and Budget standards for capturing data on race.
3. Ethnicity	The respondent is asked whether they consider themselves to be Hispanic or Latino. An additional question captures the specific Hispanic origin. The question meets the Office of Management and Budget standards for capturing data on ethnicity.
4. Gender	The respondent is asked for his/her self-conception of being male or female.
5. Current Marital Status	The respondent is asked for his/her current marital status.
6. Current Educational Attainment	The respondent is asked for his/her highest educational attainment at the current time.
7. Annual Family Income	The respondent is asked their family's total income from all sources in the last calendar year. Respondents who don't know or refuse to answer the first annual family income question are asked a series of bracketing questions to obtain the approximate amount.
8. Child-Reported Parental Educational Attainment	The respondent is asked to report on the highest level of education achieved by his/her residential parents (i.e., while growing up).
9. Impairment	Adults: The World Health Organization Disability Assessment Schedule 2.0 12-question self-administered questionnaire is a generic assessment instrument for health and disability that covers six domains: cognition, mobility, self-care, getting along, life activities, and participation. Items are rated from none, mild, moderate, severe, to extreme or cannot do, indicating difficulty over the past month. Children and Adolescents: The Columbia Impairment Scale is a 13-item self- or parent-reported instrument that measures four areas of functioning: interpersonal relations, broad psychopathological domains, functioning in job or schoolwork, and use of leisure time. Items are rated on a Likert scale ranging from 0 (no problem) to 4 (very big problem) indicating behavior at the present time.
10. Broad Psychopathology (proposed new measure)	Adults: The DSM-5 Level 1 Cross-Cutting Symptom Measure is a 23-item self-report or informant-rated questionnaire that assesses 13 domains that are important across psychiatric diagnoses. Each item inquires about how much (or how often) the individual has been bothered by the specific symptom during the past 2 weeks. Children and Adolescents: The Strengths and Difficulties Questionnaire is a parent or teacher report instrument that consists of 25 Likert-style questions that screen children for positive and negative attributes.

Core: Tier 1 measures are considered to be essential to all areas of mental health research and should be included in all National Institute of Mental Health funded human subjects research.