Supporting Statement B

The SSA-NIH Collaboration to Improve the Disability Determination Process: Calibration II & Predictive Validity of IRT-CAT Tools
NIH/CC/RMD

Date 08/11/2014

Name:	Daniel Hobbs
	Management Analyst
	Epidemiology & Biostatistics Section
	Rehabilitation Medicine, CC, NIH
Address:	10 Center Drive, Bld. 10, 1-1469
Telephone:	301.496.3817
Fax:	301.480.0415
Email:	Daniel.Hobbs@nih.gov

Table of Contents

B. STATISTICAL METHODS

B.1 Respondent Universe and Sampling Methods

A calibration study is a field study of item content and structure conducted with samples of respondents representing the intended users of the computer adaptive testing (CAT) instruments. Item pool development and subsequent item calibration are unique for each CAT tool and for the target population for which they are developed. Sample size is determined based on the statistical need to support a series of confirmatory factor analyses and to perform statistical modeling. Inadequate sample size may lead to inaccurate and unstable statistical outcomes.

(1) The Social Security Administration-National Institutes of Health- Boston University (SSA-NIH-BU) team collaborated with the subcontracted national survey firm, Westat, to establish the calibration design and sampling strategy. SSA's Office of Data Analysis will extract from SSA automated records claims submitted within the last two months and provide the following variables: claimant name, address, phone number, impairment allegations listed on the form SSA-3368BK (the Adult Disability Report). Of the 20,000 potential participants who receive a pre-notification package, it is estimated 7,800 claimants will complete screener interviews. Of those, it is estimated that 3,500 claimants will choose to participate in the study.

From this dataset, Westat will assign a geographic variable to the data to classify claimants into urban or rural categories. The sample will then be stratified by urban/rural status across the 10 national SSA office regions.

The data analysis will address the following parameters:

- Response burden
- Score precision
- Internal consistency reliability
- Score range (ie., floor or ceiling effects)
- Predictive validity

To monitor the Boston University-Health & Disability Research Institute Functional Assessment Batteries in real time, we will calculate the standardized log-likelihood statistic (*lz*) for polytomous items to test the person fit. The empirical distribution of the log-likelihood statistic is reasonably close to a standardized normal distribution, so we will calculate the percentage of respondents in which *lz* exceeded an alpha level of .05.

Response burden will be measured as the average amount of time it takes to complete instrument. A t-test will be used to assess whether the average amount of administration time between the BU-HDR FABs and other measurements is significantly different.

To illustrate the difference in precision in score range across instruments, we will calculate the average Standard Error (SE) along the entire scale continuum across different instruments. We will use the t-test to assess whether the average SE is significantly different between BU-HDR FABs and other measurements at different score ranges.

To examine internal consistency, we will use marginal reliability calculations that are specific to item response theory (IRT) which allow us to compare BU-HDR FABs with other instruments. Marginal reliabilities are similar to Cronbach's alpha coefficient used in classical measurement theory in that it is a measure of how well items within a domain relate to each other. The percentage of ceiling and flooring will be calculated in each instrument. A chi-square test will be used to test whether the percentages of ceiling or flooring are significant different between BU-HDR FABs and other instruments.

In addition to collecting data from SSA claimants, data from a normative sample of 2,000 will be collected. The normative sample data allows the research team to expand the breadth of each scale developed compared to use of claimant data alone. This will reduce ceiling effects and broaden the overall applicability of each CAT scale developed. Secondly, having calibration data from a normative sample of adults in the country provides a useful reference population against which SSA claimants can be compared. This allows SSA to better characterize their population of claimants over time.

(2) The normative national sample will be obtained using sample matching; a methodology pioneered by YouGov Polimetrix, Inc. (YGP; Palo Alto, CA) whereby samples representative of a study-appropriate target population can be constructed from large but unrepresentative pools of opt-in survey respondents. The enumeration of the target population would in traditional sampling be known as the sampling frame and would serve as the source from which the sample would be drawn. This is not the case in sample matching, which instead proceeds in two-stages.

First, a random sample is drawn from the enumeration of the target population. A simple random sample (SRS) could be drawn; but in practice, the efficiency of the procedure can be improved by using stratified sampling. YGP typically stratifies on race, gender and age, and then draws a SRS from each of the mutually exclusive and exhaustive groups formed by the simultaneous cross-classification of the population on these three attributes. The SRS from each category is combined to form the stratified target sample. If the number of respondents selected in each stratum is proportional to their frequency in the target population, then the sample is self-representing.

Conventionally, one would then attempt to contact the respondents in the target sample. However, there is no economical way of reaching most members of the target sample, as they have not provided their email addresses and many do not have a listed phone number, and those

that do, may not agree to be interviewed. Instead, for each member of the target sample, YGP will select one or more matching members from their pool of opt-in respondents. This pool has been recruited by a variety of means and currently numbers approximately 1.5 million. Of course, data drawn from this pool would not be representative of any particular population; individuals who opt-in for taking web surveys have different demographics than either the population of all internet users or the population of all adults. Rather, the matching methodology is required to produce usable samples for individual studies. Matching is done on a large set of variables available in both the population enumeration database and the opt-in panel. The purpose of the matching is to find an available respondent who is as similar as possible to the selected member of the target sample.

YGP employs a proximity matching method whereby a distance function is computed for each attribute to define the degree of "closeness" between each individual in the target sample (x) and those in the opt-in survey panel (y). Typically, the distance function is the simple absolute value of the difference, |x-y|, and the overall distance between a member of the target sample and a member of the panel is a sum of the distance functions for each attribute being used in the matching. The distance functions can be weighted and then summed if particular variables are thought to be more important for a given study. For this study, matching was done on gender, racial/ethnic background, age, education and employment status, weighted equally. We utilize standard OMB categories to capture racial/ethnic background.

YGP adjusts for anticipated non-response by selecting multiple best matches in the opt-in panel for each member of the target sample. The number of matches is determined by using a hazard model to estimate the probability that an opt-in panelist will respond by the end of the data collection period, and increasing the number of panelists matched to the member of the target sample until that response probability is >=1. Although internet use was initially concentrated in the more affluent and better-educated segments of the population, this "digital divide" has been substantially reduced such that 81 percent of the US population now has access to the internet[1]

B.2 Procedures for the Collection of Information

Westat will mail an advance mailing **(Attachment 3)** to the randomly selected sample of claimants. Information in the mailing will be drafted at a 7th grade reading level and includes an invitation letter and consent form. The advance mailing will not contain web survey log-in information. The invitation letter explains the following:

1. Study overview;

[1] "Percentage of Individuals using the Internet 2000-2012" International Telecommunications Union (Geneva), June 2013, retrieved June 4, 2014.

2. Purpose of the study;
3. Acknowledge claimants are randomly selected;
4. Claimants are under no obligation to participate in the study;
5. SSA will not know if a claimant participates in the study or not;
6. Whether the claimant takes part in the study will not affect the claimant's current disability claim or any future claims;
7. SSA will not have access to the identifiable information the claimant provides to the survey center;
8. Information collected will be used only for the purposes of this research.

Procedure:

1. Westat interviewers will follow up the advance mailing with a telephone call. During this phone call, the Westat interviewer will conduct a short screening interview **(Attachment 8)** to confirm claimant's eligibility and willingness to participate in the study. Of the 20,000 potential participants who receive a pre-notification package, it is estimated 7,800 claimants will complete screener interviews. Of those, it is estimated that 3,500 claimants will be eligible to participate in the study.

2. If the claimant is eligible, the interviewer will review the consent **(Attachment 8)** verbally over the phone, obtain verbal consent from the claimant, and provide the claimant with the option of completing the 2 new item banks (1) Daily Activities and (2) Learning and Applying Knowledge / survey 1 **(Attachment 6)** online or over the phone with an interviewer.

3. If a claimant chooses to do survey 1 online by him/herself, the interviewer will obtain the claimant's email address and send web survey log-in information and instructions to the claimant via email. If a claimant opts to do survey 1 by him/herself online and fails to do so within 4 days, Westat will send a reminder email. After 7 days, an interviewer will follow-up with them by phone and attempt to administer survey 1 over the phone.

4. If the claimant chooses to complete survey 1 over the phone with an interviewer, the interviewer may proceed with the survey immediately following the screening call or make an appointment for a time more convenient for the claimant. Interviewers will enter the web survey in the very same way the claimants would if the claimant were completing the questionnaire online by him/herself. The interviewers will launch a new window in Internet Explorer and navigate from their calling screens to the survey website where they will copy and paste in the claimants' unique access code (provided to the interviewer on the screen) and record the claimants' responses to survey questions directly into the web survey application. When the interviewer accesses the web survey, he or she will read aloud to the claimant over the phone all introductory instructions, questions and response options. The interviewer will enter the respondent's answers directly into the web survey.

5. Scheduled appointments are given a one-hour time window for the call to be made. If a study participant misses a scheduled appointment, Westat will attempt to contact the participant. If no contact, the participant will go back into the main queue with other general cases for this study.

6. During the first contact, a claimant will respond to the 2 new item banks (1) Daily Activities and (2) Learning and Applying Knowledge. We estimate survey 1 will take 1 hour, and 3,500 claimants will participate in this study component. They will receive $20 upon completion of Survey 1.

7. All claimants who complete survey 1 will then be re-contacted, approximately 10 days after the initial response, and will be administered short forms of the Physical Function (PF) and Behavioral Health (BH) FAB tools as well as FAB replenishment items (this will be referred to as Survey 2) **(Attachment 7)**. During this call, the Westat interviewer will conduct a brief screener to ensure we are speaking to the correct person, confirm their interest in participating, and provide them with the option of completing Survey 2 online or over the phone the phone with an interviewer **(Attachment 9)**.

8. If the claimant chooses to complete Survey 2 over the phone with an interviewer, the interviewer may proceed with the survey immediately following the screening call or make an appointment for a time more convenient for the claimant. If a claimant chooses to do Survey 2 online by him/herself, the interviewer will obtain the claimant's email address and send web Survey 2 log-in information to the claimant via email. If a claimant opts to do Survey 2 by him/herself online and fails to do so within 4 days, Westat will send a reminder email and after 7 days an interviewer will follow-up with the claimant by phone and attempt to administer Survey 2 over the phone.

9. We estimate survey 2 will take one hour and 3,000 claimants will participate in this component of the study. Claimants will receive an additional $30 upon completion of Survey 2.

10. If a claimant completes Survey 1 or 2 via the web, the claimant will be asked to read the consent and will be required to a check a box indicating the consent was read prior to continuing with the survey. If the claimant opts to complete the survey over the telephone with an interviewer, the consent will be collected verbally over the telephone.

11. The web survey system that Westat has developed for this project will not contain or be linked to claimant identifying information. Only Westat's unique identification number for each claimant along with his or her survey responses will be stored on Westat's system and the BU-CAT-SMS.

12. The collection of calibration data does not include use of paper questionnaires.

13. Claimants who choose to participate in the study via telephone interview will have their responses entered directly into the survey application by the Westat interviewer.

14. Of the approximately 100 Westat personnel who have received SSA suitability clearances to participate on the project, approximately 30 staff members are designated as working in the Telephone Research Center (this includes data collectors, Supervisors, team leaders, inbound personnel, etc). From the pool of data collectors with clearance, fewer than 50 of them may actually work on the project. Data collectors will enter survey response data directly into the BU-CAT SMS.

15. To achieve the desired sample size of 3,000 with completed Survey 1 and 2 responses, we anticipate the collection of approximately 200-250 surveys per week.

16. Each study participant will be a single record (transaction).

17. Each participant will be assigned a unique ID. Data will be associated with the unique participant ID, not the SSN. SSNs will not be requested from SSA or the claimant.

The normative national sample will be obtained using sample matching; a methodology pioneered by YouGov Polimetrix, Inc. (YGP; Palo Alto, CA) whereby samples representative of a study-appropriate target population can be constructed from large but unrepresentative pools of opt-in survey respondents.

B.2.1. Quality Control

The contractors for this study will establish and maintain quality control procedures to ensure standardization, and high standards of data collection and data processing. The contractor will maintain a log of all decisions that affect sample enrollment and data collection. The contractor will monitor response rates and completeness of acquired data.

B.3 Methods to Maximize Response Rates and Deal with Nonresponse

During the first contact of the calibration study, a claimant will respond to the 2 new item banks (1) *Daily Activities* and (2) *Learning and Applying Knowledge*. We estimate this will take no more than 1 hour, and we will provide a $20.00 incentive for voluntary participation in this study component. These individuals will then be re-contacted, approximately 1-2 weeks after the initial response, and will be administered the short-form *Physical Function* and *Behavioral Health* FABs, as well as replenishment items. We estimate this will take no more than 60

minutes (1 hour). We will provide an additional $30.00 incentive for voluntary participation in this component of the study.

YPG's goal is to provide a small thank you for a respondent's time, but not an incentive that might make survey response a financial transaction. The average survey incentive of 500 points cashes out at 50 cents, although it's not redeemable until respondents reach certain thresholds. Respondents who complete the instruments required for this normative population portion of the study will earn approximately 1,500 points. This level of remuneration is within the industry standard range, and has been approved by the appropriate Institutional Review Board.

Based on the Westat research team's previous experiences with Calibration studies, they have developed very structured methods to ensure the maximum possible response rate. These methods include scheduling of follow up (Survey 2) at the conclusion of the initial contact, providing multiple modes of administration (phone interviewer or email weblink), and the option for courtesy reminders.

The calculated response rate will be determined as follows:

Calculating response rates for Survey 1:
Total Survey 1 completes/(Total Survey 1 sample called – all ineligibles)

Calculating response rates for Survey 2:
Total Survey 2 completes/(Total Survey 1 completes called)**

** Note that all Survey 1 completers are eligible for Survey 2; therefore, there are no ineligibles.

B.4 Test of Procedures or Methods to be Undertaken

Preliminary testing of the workflow throughout Westat's survey management system, programmed to collect calibration study data, was enhanced with the receipt of dummy data in advance of data collection. The Office of Data Analysis (ODA) pulled dummy data for the Office of Disability Policy (ODP). ODA sent a file to ODP containing data for which ODA believed that they effectively removed PII. ODP requested the ODP and ORDP Security Officers review the information and approved the release of the dummy data to Westat.

B.5 Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

Leighton Chan, MD, MPH Chief, Rehabilitation Medicine Clinical Research Center National Institutes of Health Department of Health and Human Services 10 Center Drive Bethesda, MD 20892-7362 Phone: (301) 496-4300 Fax: (301) 480-4970 Email: chanle@cc.nih.gov	**Elizabeth Rasch, PT, PhD** Chief, Epidemiology & Biostatistics Rehabilitation Medicine Department Clinical Research Center National Institutes of Health 6100 Executive Boulevard Bethesda, MD 20892-7344 Phone: 301-496-3817 Fax: 301-480-7515 Email: rasche@cc.nih.gov
Alan Jette, PT, PhD Director Health & Disability Research Institute Boston University School of Public Health 715 Albany St. Boston, MA 02118 Phone: (617) 638-1985 Email: ajette@bu.edu	**Diane Brandt, PT, MS, PhD** Protocol Manager Epidemiology & Biostatistics Section Rehabilitation Medicine Department Clinical Research Center, National Institutes of Health Phone: 301-496-3817 Email: brandtd@mail.nih.gov
Peter Smith, JD Project Officer Social Security Administration ORDP/ODP/OCADO 6401 Security Blvd, 4531 Annex Baltimore, MD 21235 410-965-2239 (phone) 410-597-0280 (fax) Email: Peter.Smith@ssa.gov	**Beth Marfeo, PhD, MPH** Research Instructor Health Policy & Management Health & Disability Research Institute BU School of Public Health Phone: (617) 638-1987 Email: bmarfeo@bu.edu
Minh Huynh, PhD Labor Economist, Analytic Lead Epidemiology & Biostatistics Section Rehabilitation Medicine Department Clinical Research Center, National Institutes of Health Phone: 301-496-3817 Email: huynhm@cc.nih.gov	**Mary Slavin, MS, PhD** Research Professor Health Policy & Management Health & Disability Research Institute BU School of Public Health Phone: (617) 638-1987 Email: mslavin@bu.edu
Kara Bogusz, MA Project Administrator Health Policy & Management Health & Disability Research Institute BU School of Public Health	**Erika Bonilla, MS** Research Coordinator Westat 1600 Research Blvd. Rockville, MD 20850

Phone: (617) 638-1987 Email: kbogusz@bu.edu	Phone: (800) WESTAT1 (ext. 4879)
Christine McDonough, MS, PhD Research Assistant Professor Health Policy & Management Health & Disability Research Institute BU School of Public Health Phone: (617) 638-1987 Email: cmcdonough@bu.edu	

List of Attachments

Attachment 2 – Pre-notification letter with consent form

Attachment 4 – Complete Item Bank for Daily Activities & Learning and Applying Knowledge Functional Assessment Batteries (Survey 1)

Attachment 5 – Short-form versions of Physical Function & Behavioral Health Functional Assessment Batteries including Replenishment Items (Survey 2)

Attachment 8 – Screening Interview Script Survey 1

Attachment 9 – Screening Interview Script Survey 2