NCS Formative Research Template for OIRA Clearance

TO BE COMPLETED BY FIELD CONTRACTOR:

LOI #:   LOI [ ] [ ] [ ]

Title of Formative Research: Assessment of Executive Function for the National Children's Study
Participating Institutions: University of Minnesota, Northwestern University, Delve
Marketing Research

SME: **Gitanjali Taneja**
COR: **Davyd Chung**

Purpose of the Study:
This project was designed to develop robust and brief measures of executive function (EF) for the NCS suitable for preschool-aged children (ages 2.5 to 5.5). The purpose of the current study is to refine and validate the measures, develop materials required to train administrators who do not have a professional background in child development or testing, and publish results of this work. Please see Attach. 21 IRB Orig Application (pp. 7-8) for more information.

Benefit to NCS:
It is widely recognized that EF skills are important for many aspects of human health and development, including physical and mental health and educational success of children and adults.[1] Thus, it is critical to assess EF skills in the National Children's Study in a manner that is sensitive for young children. Currently, available measures (used in the Vanguard Study) were not developed for the assessment of diverse cohorts of children, especially in regard to psychosocial disadvantage and ethnic or cultural diversity. It is important that the measures adopted by the NCS be efficient and also broadly suitable for the full range of families that will be included in the NCS.

One of the central aims of this project is to examine the reliability and validity of the revised EF measures that were explicitly revised to expand the range of the measures for more diverse children, specifically including disadvantaged children, including children from homeless and very low-income families. It is crucial that the National Children's Study use measures of cognitive development that work well with a diverse range of children. Growing evidence indicates that children from low SES backgrounds have higher risk for EF problems and also that EF skills are related to resilience among high risk children, including homeless children.[2] We plan to test the validity and reliability of the new measures in relation to income and education level. In order to evaluate whether income and education levels in the family relate to the validity or reliability of the new measures, we need to include measures of both.

This project is thus designed to adapt existing EF measures in order to investigate whether they (1) have good evidence of validity and reliability for middle class or more advantaged samples, (2) demonstrate similar robustness when used with less advantaged or with culturally diverse populations, and (3) are sensitive to capturing useful variance in the youngest participants. Three key measures will be adapted for preschool-aged children (ages 2.5 to 5.5):
   (1) Flanker (Attach. 3), a computer- administered test from the NIH Toolbox assessment battery that was developed to measure EF for individuals spanning ages 3 to 85,
   (2) Dimensional Change Card Sort (Attach. 2), also a computer- administered test from the NIH Toolbox assessment battery that was developed to measure EF for individuals spanning ages 3 to 85, and
   (3) Children's Behavior Questionnaire with EF extension (CBQ-VSF, Attach. 10).

One of the main goals for this study is validation of these measures with respect to time burden, usability, reliability, and construct validity in diverse families with preschool-aged children ages 2.5 to 5.5. Thus, additional measures used for validation purposes are also included: WPPSI-IV Block Design (Attach. 7), Woodcock-Johnson-III Applied Problems, Letter-Word Identification (Attach. 8), Peg tapping

---

[1] Blair & Razza, 2007; Buckner, 2008; Carlson & Zelazo, in press; Obradovic, 2010.

[2] Balri & Raver, 2012; Karatoreos & McEwen, 2013; Masten, 2012; Masten et al. 2012; Obradovic, 2010; Raver, Blair, & Willoughby, 2013.

(Attach. 6), and the rest of the NIH Toolbox cognition battery, which has been adapted for use with an iPad.

Since EF is very strongly linked to mental and physical health outcomes, developing sensitive and valid measures of EF in early childhood is of paramount importance, and therefore, a short and financially feasible study will be of immense value to the NCS in providing a research platform for future study.

Study Design:
Data will be collected in three sites in different regions of the country: Minneapolis, MN, Phoenix, AZ, and Philadelphia, PA. Data collection will be conducted through the University of Minnesota (in Minneapolis only) and Delve Marketing Research (at all 3 sites). Delve was the firm that collected the national norming study data for the original NIH Toolbox measures. Northwestern University will work with the University of Minnesota team to coordinate the study, and the University of Minnesota team will collect data for the subsample of disadvantaged children.

The University of Minnesota will supervise training and data scoring as required to complete the study. The Minnesota team will be responsible for training the staff both at the University of Minnesota and Delve on the consent procedures as well as test administration (at the Delve Minneapolis site and at the University of Minnesota). Delve personnel will be added to the personnel for this project at a later time (after OMB approval, funding, and hiring, but before any data collection for this project).

Families will be recruited with strategies adapted to each context, using established and IRB-approved methods that have worked well in each of these contexts see the Methods of Recruiting section below. Once consent forms (Attach. 16-19) have been signed (and children have orally assented)(Child Assent Form, Attach. 26), the parent and child will begin assessments simultaneously, either in the same room or in separate rooms near each other.

<u>**Session 1 (Test)**</u>
Children at all sites will be tested individually in a single session of 50 to 60 minutes, with parents completing the parent-report measures simultaneously. The following measures will be administered (see Attach. 1 - 13 for more information about each measure):

Child: (53 min)
Peg Tapping (3 min)
Touch Screen Tutorial (3 min)
NIH Toolbox Picture Vocabulary Test (5 min)
NIH Toolbox Flanker Inhibitory Control & Attention Test (7 min)
NIH Toolbox Dimensional Change Card Sort Test (7 min)
NIH Toolbox Picture Sequence Memory Test (5 min)
WPPSI-IV Block Design (8 min)
Woodcock-Johnson-III Tests of Achievement (Letter-Word Identification-7 min, Applied Problems-8 min)
Parent: (35 min)
Family Demographic Questionnaire (7 min)
Children's Behavior Questionnaire Very Short Form with EF extension (CBQ-VSF+EF) (10 min)
Ages and Stages Social Emotional Questionnaire (10 min)
Quick Assessment of Behavior in Children (Q-ABC) (3 min)
Strengths and Difficulties Questionnaire (5 min)

<u>**Session 2 (Retest)**</u>
A subset of children and their parents will be asked to come back 1-2 weeks later for a follow-up session. The parent and child will be tested simultaneously with a child session of about 30 minutes and the parent session of 10 to 15 minutes. The reliability of the following measures will be tested:

Child: (30 min)
Peg-tapping (3 min)
Touchscreen tutorial (3 min)
NIH Toolbox Picture Vocabulary Test (5 min)

NIH Toolbox Flanker Inhibitory Control and Attention Test-Downward Extension (7 min)
NIH Toolbox Dimensional Change Card Sort Test-Downward Extension (7 min)
NIH Toolbox Picture Sequence Memory Test (5 min)
Parent: (13 min)
Children's Behavior Questionnaire Very Short Form with EF extension (CBQ-VSF+EF) (10 min)
Quick Assessment of Behavior in Children (Q-ABC) (3 min)

Analyses will focus on information pertinent to administration time and psychometric data on reliability and validity, including the following analyses. The University of Minnesota team will provide descriptive data on the administration time and scores for all measures, including means, SDs, range, overall and broken down by age and gender. Reliability will be analyzed, including (as appropriate to the measure), internal consistency (alpha) and test-retest reliability, overall and broken down by age and gender groups for finer examination. Multivariate analyses (e.g., correlations, regressions) will be conducted to analyze the convergent, discriminant, and construct validity, for example providing tables and figures related to correlations of new measures with age, each other, other measures of EF, traditional IQ measures, and criterion measures of school readiness (literacy and numeracy) that have established predictive validity, both overall and with age covaried. Analyses examining relation of child scores to family SES, income, and parent education level also will be conducted. A technical report and at least one paper suitable for publication will also be submitted at the conclusion of the contract period.

To determine whether the obtained estimates of reliability and validity are sufficiently high to recommend including these measures into the National Children's Study, we will compare them to estimates obtained from the NIH Toolbox Validation study for the cognition measures (Zelazo & Bauer, 2013) and those reported in the technical manuals of the standardized gold-standard measures. We will additionally refer to measures of reliability and validity that were obtained in studies conducted by the MN team with the same disadvantaged population that will be recruited for the proposed study. [3, 4]

Across the NIH Toolbox cognition measures, intraclass correlation coefficients (ICC; estimate of test-retest reliability) less than .4 were considered poor, $.4 <= ICC < .75$ were considered adequate, and $ICC >= .75$ were considered excellent. The obtained ICC estimates for children ages 3-6 were .73 for Flanker and . 87 for DCCS (in the *excellent* range). Convergent validity was assessed with correlations between each NIH Toolbox measure and a well-established validation measure of the same construct. Across measures, $r < .3$ were considered poor, $.3 <= r < .6$ were considered adequate, and $r >= .6$ were considered excellent. The obtained construct reliability estimates for children ages 3-6 were .63 for Flanker and .68 for DCCS, when compared to the gold standard of the WPPSI Block Design (again in the *excellent* range).

For WPPSI-Block design, administered to children 2.5-5.5 years of age, the test-retest reliability scores (measurements separated between 14 and 50 days) measured with the Pearson product-moment correlation coefficients were between r=.69 and .81, depending on the specific age bracket. For the Woodcock-Johnson measures, the test-retest reliability measures (measurement separated by one year) for children 4-7 years old (data for younger children are not available) are r=.92 for both Letter-word identification and Applied problems.

With the disadvantaged population, the previous studies conducted by the MN researchers yielded validity estimates for the Flanker, DCCS and Peg Tapping between .3 and .48, with test-retest reliability estimates of .4 for DCCS, .5 for Flanker and .8 for Peg Tapping. Thus, the obtained reliability and validity scores fell into the "adequate" range even for the disadvantaged populations.

For the "typical" population, we expect levels of reliability and validity to again fall into the *excellent* category (based on the NIH Toolbox Validation and the technical manuals for the gold-standard

[3] Masten, A.S., Herbers, J.E., Desjardins, C.D., Cutuli, J.J., McCormick, C.M., Sapienza, J.K., Long, J.D., & Zelazo, P.D. (2012). Executive Function Skills and School Success in Young Children Experiencing Homelessness. Educational Researcher, 41, 375-384.

[4] Obradjovic, J. (2010). Effortful control and adaptive functioning of homeless children: Variable-focused and person-focused analyses. Journal of Applied Developmental Psychology, 31, 109-117.

validation measures). We expect the estimates to be somewhat lower for the disadvantaged group, but to nevertheless fall into at least the "adequate" range of scores.

Training of Staff
A two-day "in-person" training of all staff (both at the University of Minnesota and Delve) will be implemented. Trainees will pass a rigorous examination process to test proficiency with all the measures proposed for this study, in order to obtain certification that will enable them to collect data.

Trainees would be expected to have experience testing children; this will be one of the criteria for hiring. They will review instructions and demonstration videos for some measures in advance of the training sessions. The first day of training will focus on child measures: orientation, instructions for the battery, demonstrations, and initial practice. Further practice will be available later in the day and/or evening. On the second day, each assessor will demonstrate the administration of tests and receive individualized feedback to improve fidelity of administration of child tasks and also review the parent measures (the training of the parent measures can be briefer, since these are straightforward questionnaires). Each day will include some discussion of "what to do" when common challenges arise and how to handle exceptional cases. Training will be supervised by the Minnesota team members. If a trainee does not meet the criteria for fidelity during the training process, then additional practice will be set up for later observations until the trainee meets the expected administration quality.

All measures proposed to be used in this study have objective criteria and do not require examiner's judgment in scoring the tests. In addition, the NIH Toolbox iPad measures have been designed to be maximally intuitive and straightforward to administer, such that no judgment from the examiner is required on any items in order to successfully administer the tests. All instructions will be printed on the screen for ease of administration.

All Delve staff will be trained on site at the Minnesota Delve office. If further observation is needed, that will occur at each testing site. Minnesota staff would be trained at the University as per their usual protocol.

In addition, the Minnesota and/or Northwestern University staff will visit periodically and observe assessors at each site to check on fidelity of administration and compliance with all security measures. The specific protocol will be as follows:
1. A site visit by NU researchers for each of the 3 Delve Sites will be scheduled early in the data collection process.
2. NU and/or MN researchers will engage in occasional observation of examiners (once or twice for each examiner) via in-person, Skype, or video recordings.
3. Weekly check-in with examiners to address any questions/issues will be implemented.
4. Thorough checks of the Behavior Rating Sheets, which is a 10-item report completed by the examiner about how well the session with the child flowed, and other completed paper documents (that have been scanned and uploaded) will be conducted 2 times a week by a MN coordinator.
5. Slow rollout of data collection will be implemented in the first few weeks, with particularly thorough checks of the Behavior Rating Sheets and checking in with the examiners to determine if anything is unclear.
6. NU will ensure that data could be examined as a function of site/examiner to check if some sites/examiners have different distributions of scores.

Pilot Study
The University of Minnesota conducted a pilot study with N=9 child/parent dyads as part of the NCS formative project.

Sample demographics: Children ranged in age from 3 years 6 months to 5 years 8 months (4 girls, 5 boys). Parents ranged in age from 23 to 44, with the average age about 30 years old. In six families, the parent participant was the biological mother, in one the adoptive mother, and in another the biological father. These families were very diverse, with most being African American, and some having Caucasian, biracial, Native American, and Hmong family members. Parents' highest levels of education ranged from partial high school to an associate's degree. These participants came from two of the

planned sites where they propose to collect the new low-income sample (a homeless shelter and a community preschool for very disadvantaged children).

Methods: Parents, children, and teachers were asked to complete the following measures:

Parent		Child	
Family Information Quex	7 min	Touch Screen Tutorial	1-2 min
Touch Screen Tutorial	2min	Flanker	~5 min
Flanker	5 min	DCCS	~5 min
DCCS	5 min	Stanford Binet	15 min
CBQ	10 min	Bracken*	10 min
Ages & Stages	10-15 min	WJ-Applied Problems*	5 min
Q-ABC	2 min	WJ-Letter Word Identification*	5 min
		TOPEL*	15 min

These school readiness assessments were piloted to assess time burden and appeal of measures. Eight children did the Bracken, 8 completed the WJ Applied Problems, 5 did WJ Letter/ Word Recognition, and 2 received TOPEL.

Parents and children alike showed enthusiasm for the measures. At the end of each session, parents were asked about their favorite and least favorite parts of the sessions. The majority of parents indicated that they liked getting to spend time thinking about and telling others about their child, whereas others expressed that they liked the computer games.

Results: Data from DCCS and Flanker (the two main measures of executive function) were shown to correlate with each other (r=.75) and improve with age (r=.73 for DCCS and r=.89 for Flanker). Scores on these two measures also correlated with raw scores on a nonverbal IQ measure; However, performance on the extended DCCS and Flanker measures were not related to scaled scores (age-normed scores) on the nonverbal IQ measure, suggesting that the EF measures captured unique variance beyond any covariance with traditional IQ measures. Results further suggested that the associated time burden was appropriate, and the tasks appealed to the participants. Results were used to further refine the EF downward extension ("dext") paradigms proposed here, including streamlining and simplifying instructions. Findings were presented by the University of Minnesota team at the NCS Research Day in August 2011, both as an oral presentation, and in a poster.

Target Respondents:
Study subjects will consist of: 1) children between 2.5 and 5.5 years of age and their primary caregivers, 2) native speakers of either English or Spanish, and 3) children without a parent-reported disability that would preclude fully participating (such as a pervasive developmental disability, blindness, or deafness).

Children and parents who do not speak English and with a parent-reported disability that would preclude full participation (such as a pervasive developmental disability, blindness, or deafness) will not be included in the study.

The number of participants to be recruited is 396 children (approx.50% female, 50% male) ages 2.5 to 5.5 and 396 adult parents/guardians (80% female, 20% male). Thus, the total number of participants to be recruited is 792. They will be recruited though either the University of Minnesota or Delve Marketing Research. They will be non-NCS participants.

A random subsample of children and their parents (N=144 families) will be retested after 1 to 2 weeks to test reliability of measures.

Sample Size Calculation:
The total proposed child *N* for testing is 360, with a test-retest sample of 144. In addition, 360 parents/guardians will also participate. In order to achieve these numbers, we have included a10% larger sample of child-parent dyads given the possibility of children deciding to stop mid-assessment or being unable to continue for any reason (e.g., tired or becoming ill); thus, IRB permission has been requested and approved for 396 children and 396 parents/guardians. We will use a convenience sample. The retest sample size was chosen based on the same proportion of participants that were asked to participate in a retest session for the NIH Toolbox Validation study (Zelazo & Bauer, 2013). This proportion was shown to be an appropriate compromise between minimizing burden (such that only 40% of participants are asked to come back) while obtaining a sufficiently large sample to determine rest-retest reliability estimates for each age year.

The University of Minnesota team will collect data as planned and approved for very disadvantaged children
residing in shelters or attending community preschools for disadvantaged children. The goal is to assess 60 children (60 parents/guardians) with a random subset of approximately 30 retested for reliability.

Delve will collect data for 240 children (and 240 parents/guardians) in English and 60 children in Spanish (and their 60 parents/guardians) at 3 locations, divided approximately equally across sites. A random subsample of children will be retested after 1 to 2 weeks to test reliability, including 90 in English and 24 in Spanish (30/8 at each site). See Table 2 below for the retest target sample.

The target sample breakout by age and demographic variable is provided in the Table 1 below. We expect that the children will be about equally divided by sex. We expect the parents/guardians to be primarily female. Sampling recruitment will aim for balanced age and sex within six 6-month age brackets: 2.5 (older than 2.5) through 5.0-5.4 (younger than 5.5).

The sample size (360 total child participants) was determined to be appropriate based both on a formal power analysis and previous work. Power analysis indicated that the proposed sample size of N=360 would provide sufficient power (over 90%) to detect whether a correlation between a Dext measure of EF and the parent-reported EF of 0.20 is significantly different from 0 (parent reports of EF and child test scores typically fall in this range). In addition, the stratification by age and gender of N= 30 Males and 30 Females for each six-month bracket enables examining the effectiveness with which EFs are measured as a function of each year of age and each gender.

We are including both the "typical" and the disadvantaged participants in the overall sample, because we aim to ensure that the proposed measures of executive function are appropriate for a wide range of participants. The disadvantaged sample (N=60) will be equally divided into the six age bands (N=10; 5 male and 5 female) for each six-month age bracket. As described on p. 3, we expect to see lower estimates of both reliability and validity for the highly disadvantaged sample (while still falling in the "adequate" range), with decreased group means and increased variability relative to the typical population (recruited by Delve).

In addition, previous research helped to inform the proposed sample size. A comparable sample size (50 Males and 50 Females for each 1 year bracket) was successfully used in the NIH Toolbox Norming study (from which most of the measures proposed here will be adapted). The ultimate goal of the Toolbox Norming project was to provide normative reference values that would be acceptable to the epidemiological and clinical research community. This sample size per cell was selected based on the precedent of other norming studies, in particular the Wechsler Adult Intelligence Scale. Therefore, based both on the formal power analysis and previous research, the sample size of 360 children is deemed to be appropriate for this Formative Research study.

It should be noted that the sample of homeless/highly mobile families from low-income backgrounds will include higher rates of minority children and parents, and lower education levels on average (reflecting the demographics of this population in Minneapolis and many other urban centers across the country). The goal of this sample is to explicitly ensure that the measures show reliability and validity and work well in the field with diverse and disadvantaged children. This subsample will be well-distributed by age

and balanced by gender, although a somewhat greater degree of flexibility in accrual will be allowed to facilitate participation of families residing in emergency shelters. Participation rates of this population have been excellent in previous studies by the Minnesota team.[5, 6]

Table 1. Target Samples by Age and Demographic Variables

Child Sample						
Age*	**English**		**Spanish**			**Race/ Ethnicity****	**Parent Educational Attainment**
	Male	**Female**	**Male**	**Female**			
2.5-3.0	25	25	5	5	**Targets for age groups:**	16% H, 20% B, 10% O	25% < HS, 25% HS, 50% some college+
3.0-3.5	25	25	5	5		16% H, 20% B, 10% O	25% < HS, 25% HS, 50% some college+
3.5-4.0	25	25	5	5		16% H, 20% B, 10% O	25% < HS, 25% HS, 50% some college+
4.0-4.5	25	25	5	5		16% H, 20% B, 10% O	25% < HS, 25% HS, 50% some college+
4.5-5.0	25	25	5	5		16% H, 20% B, 10% O	25% < HS, 25% HS, 50% some college+
5.0-5.5	25	25	5	5		16% H, 20% B, 10% O	25% < HS, 25% HS, 50% some college+
Total	150	150	30	30			

*Age is as of initial testing date.
**Race/ethnicity targets apply to English-speaking group only and are +/-5%. It is expected that the Spanish-speaking group will be almost entirely Hispanic.
H=Hispanic, B=Black, O=Other, includes Asian, American Indian, Alaska Native, Native Hawaiian or Other Pacific Islander and individuals with 2 or more races.

Table 2. Target Sample for Retest by Age.

Child Retest Sample				
Age	**English**		**Spanish**	
	Male	**Female**	**Male**	**Female**
2.5-3.0	10	10	2	2
3.0-3.5	10	10	2	2
3.5-4.0	10	10	2	2
4.0-4.5	10	10	2	2
4.5-5.0	10	10	2	2
5.0-5.5	10	10	2	2

[5] Masten, A.S., Herbers, J.E., Desjardins, C.D., Cutuli, J.J., McCormick, C.M., Sapienza, J.K., Long, J.D., & Zelazo, P.D. (2012). Executive Function Skills and School Success in Young Children Experiencing Homelessness. Educational Researcher, 41, 375-384.

[6] Obradjovic, J. (2010). Effortful control and adaptive functioning of homeless children: Variable-focused and person-focused analyses. Journal of Applied Developmental Psychology, 31, 109-117.

Total	60	60	12	12

Method of Recruiting:
University of Minnesota
Families attending programs operated by The Family Partnership (formerly called Reuben Lindh Family Services), who operate several preschools, and families with children in the early childhood programs at People Serving People, which is a shelter for homeless families, will be asked to participate. The Family Partnership and People Serving People serve low-income families from culturally diverse populations. The goal is to assess 60 children (60 parents/guardians) with a random subset of approximately 30 retested for reliability of diverse backgrounds, including a high proportion of disadvantaged families.

A letter of invitation (Letter of Invitation, Attach. 14) will be given to parents at the Family Partnership by staff or delivered to the rooms or mailboxes of eligible families at People Serving People. Announcements may be made at community meetings. All recruitment materials will indicate how to contact the study staff for more information or to set up an appointment. In the experience of the University of Minnesota, families also tell each other about research opportunities, and thus we would expect some parents to approach study staff directly at the community sites. These methods have worked very well in past research in the same or similar community sites. The Minnesota team will train the staff both at Delve and at the University of Minnesota on the consent procedures as well as test administration (at the Delve Minneapolis site).

In the preschool settings, staff that routinely meet with families will inform eligible families about the study and distribute letters through routine channels at that site (e.g., mailboxes, meetings). Ann Masten, PhD at the University of Minnesota has conducted research with families from People Serving People for over 20 years. The research staff at the University of Minnesota have been collecting data for other studies of EF in this shelter collaboratively in recent years and the shelter provides dedicated space for their research. The Family Partnership (Reuben Lindh) has newly remodeled space for their preschool programs at their main site with ample room to accommodate this project.

Delve Marketing Research
The rest of the children will be recruited by Delve Marketing Research. Recruitment methods will be identical across the three sites (Minneapolis, MN, Phoenix, AZ, and Philadelphia, PA). Delve will randomly select participants from its database of potential participants. Delve uses an "opt-in" database from sources such as booths at community engagements, newspaper ads, Mommy-n-Me groups, social media, and word of mouth. These families have previously agreed to be contacted for future research studies, and the databases of its contractors, which, similarly, contain contact details of subjects interested in taking part in research. Hospital/clinic staff will not be involved. Delve staff will conduct the consent process and data collection.

Both Delve and University of Minnesota will contact potential participants by telephone and ask a series of questions to confirm their eligibility (speaking English or Spanish fluently and having a child between 2.5 and 5.5 years of age). Once the eligibility is confirmed, an appointment will be scheduled at the site office for testing. A Delve technician will also confirm the participant's mailing address and inform them that a packet – a consent form and directions to the testing site – will be mailed to that address shortly. When a child is selected, the telephone recruitment screening interview (Recruitment Screening Interview, Attach. 15) will be conducted with a parent or legal guardian. The consent form(s) will be mailed to potential participants as a source of additional information about the study. Participants will be formally consented by a trained Delve technician once they arrive at the testing site.

Privacy:
Study locations will abide by the terms of their Data Use Agreement, which should reference all formative research efforts involving the collection or management of NCS restricted-use data. All participating Study locations will have approved Data Use Agreements and Security Plans prior to launch.

Northwestern University will manage FISMA compliance in terms of programming computers used for data collection and maintaining all security, servers, and processes for uploading data daily from the computers used in data collection. The data collected on iPads will be wiped as soon as the upload occurs.

Consent forms will be retained by Delve in securely locked files in locked offices until the study is completed and then will be destroyed. Keys linking people or contact information to ID numbers will be maintained securely by each test site (in encrypted and password-protected computer files or locked regular files in locked offices, and separate from any data locations) until the study is completed and then the keys will destroyed.

In the event of an audit or other necessary processes concerning consent forms by the University of Minnesota IRB, an authorized supervisor from Delve would bring any needed consent forms to Minneapolis.

Paper measures will be scored and entered on databases in a secure FISMA environment on a secure server. The FISMA compliant server space will be provided by Northwestern University in their role as contractor. After initial entry, original forms will be scanned (identified only by ID numbers) into secure files on the same servers in order to verify scoring and data entry from distant sites. Data will be entered a second time by University of Minnesota staff with access to the scanned data on the secure server. Once these data are verified, the originals will be destroyed at each site.

For Minneapolis data, paper forms from Delve will be transported by hand to the University of Minnesota offices for storage in secure and locked files until they are verified.

For data collected by the University team in Minneapolis, paper forms will be stored in secure and locked files and double-entered into the secure database. Once the data are verified, the originals will be destroyed.

IRB Approval: Local IRB clearance for this activity has been obtained by the participating field contractor, Northwestern University (Northwestern Univ IRB Letter, Attach. 25), as well as the subcontractors University of Minnesota (Attach. 20-23, 27) and Delve (Delve IRB Letter, Attach. 24). Please see the IRB protocols and approval letters.

Incentives: As a token of appreciation for their participation, parent-child dyads will receive a maximum incentive of $75. This will be broken down as follows: $25 monetary incentive in the form of a debit or gift card given to the parent; $25 monetary incentive in the form of a debit or gift card OR a non-monetary incentive valued at $25 or less to thank the child (but given to the parent, if in the form of debit/gift card), and $25 in the form of a debit or gift card given to the parents to compensate for travel expenses IF they traveled for testing. This payment structure is consistent with previous OMB-approved studies that our group had been involved with, such as the NIH Toolbox Norming study,[7,8] where parent-child dyads received total payments of $90 ($125 if they lived in the LA area).

Only the participants who travel to the test site will receive the $25 travel allowance. Delve families will travel to central test sites. However, families in the emergency shelter would be tested on site in testing space dedicated for use by the University of Minnesota team. Parents from the Family Partnership are likely to be tested at their preschool community site, some traveling on their own, and some needing transportation, and thus may receive a monetary incentive of $50 or $25 (excluding the $25 or a toy for the child). Most of the disadvantaged families living in shelters will not be traveling to participate in this study, and will, therefore, not be receiving the $25 travel allowance incentive. Thus, only 83% (300 out

---

[7] Akshoomoff, N., Newman, E., Thompson, W. K., McCabe, C., Bloss, C. S., Chang, L., ... & Jernigan, T. L. (2014). The NIH Toolbox Cognition Battery: results from a large normative developmental sample (PING). *Neuropsychology*, *28*(1), 1.

[8] Beaumont, J. L., Havlik, R., Cook, K. F., Hays, R. D., Wallner-Allen, K., Korper, S. P., ... & Gershon, R. (2013). Norming plans for the NIH Toolbox. *Neurology*, *80*(11 Supplement 3), S87-S92.

360) of participants are expected to travel for participating in this study and thus be provided an incentive for travel expenses. Families without their own mode of transportation will need to use taxis to arrive to testing sites; thus, the $25 travel incentive amount is reasonable. To be equitable to all participants, everyone will be receiving the same amount ($25) if they are able to travel to the testing site. Parents/guardians will be asked to sign a receipt when they receive a monetary incentive. Delve routinely uses debit cards as monetary incentive. The University of Minnesota routinely uses Target or Wal-Mart cards in similar studies, which have worked well with disadvantaged families.

As a token of appreciation for their participation, children recruited by Delve will be offered a monetary incentive of $25 (debit card given to the parent) by Delve (this is their standard monetary incentive for children). Children recruited by the University of Minnesota will receive non-monetary incentives (i.e., toys, books, etc.) valued at $25 or less (the standard approach of the faculty researchers and for research in the same settings).

Sensitive Questions: The study has been designed to sample from three distinct groups - normative children from 3 communities sampled by Delve, speaking (1) English or (2) Spanish, and children from very disadvantaged families sampled by the University of Minnesota. In order to compare and contrast the findings for the three samples, with respect to reliability and validity, it will be important to know how the samples differ in parent income and education level. Thus, some of the questions necessarily deal with sensitive information regarding income, education level, and housing situation (homelessness).

We note that the University of Minnesota researchers have used a similar demographic questionnaire in the past, and they report no objection from previous participants regarding the sensitive items. [9],[10] In the consent, questionnaire, and during administration, they make it clear that all of the items are optional to answer. In addition, the items that deal with family income have a "prefer not to answer" option. Thus, every effort is made to minimize the potentially intrusive nature of the sensitive questions. Understanding the participants' socio-economic background is critical for evaluating whether the EF measures proposed to be used in the NCS are appropriate for participants from a wide range of backgrounds. The growing field of study of the effects of SES and education on many life outcomes suggests that measures of SES and education do have high construct validity.

Proposed Project Schedule: We will begin this project upon receipt of all regulatory approvals.

Data Collection Burden:

Estimates of Annual Hour Burden – Assessment of Executive Function for the NCS (LOI2-QUEX-6) A description of each session, including the total number of participants, timing, and activities, is provided below.

Session 1: The total proposed children for testing is 360, as well 360 parents/guardians, for a total of 720 participants for Session 1. After the consent forms have been signed (and children have orally assented), the parent and child will begin Session 1 simultaneously, either in the same room or in separate rooms near each other. Children at all sites will be tested individually in a single session of about 53 minutes, with parents completing the parent-report measures simultaneously in an session of about 35 minutes.

Session 2: A subset of children and their parents will be asked to come back 1-2 weeks later for a Session 2. The total proposed children for the test-retest sample is 144, as well as 144 parents/guardians, for a total of 288 participants in Session 2. After the consent forms have been signed

---

[9] Carlson, S. M., White, R. E., & Davis-Unger, A. C. (2014). Evidence for a relation between executive function and pretense representation in preschool children. *Cognitive Development*, *29*, 1-16.

[10] Harms, M., Zayas, V., Meltzoff, A. N., and Carlson, S. M. (2014). Stability of executive function and predictions to adaptive behavior from middle childhood to pre-adolescence. *Frontiers in Psychology: Developmental Psychology*. DOI: 10.3389/fpsyg.2014.00331

(and children have orally assented), the parent and child will begin Session 2 simultaneously with a child session of about 30 minutes and the parent session of 13 minutes.

Data Collection Activity	Type of Respondent	Estimated Number of Respondents	Estimated Number of Responses per Respondent	Average Burden Hours Per Response	Estimated Total Annual Burden Hours
Recruitment Screening Interview *(Attach. 15)*	Parents	396	1	10/60	66
Informed Consent I *(Parent-Attach. 17 or 19; Child-Attach. 26))*	Parent & Child	396	1	13/60	86
Session I (Test)* *(Parent-Attach. 9-13; Child-Attach. 1-8)*	Parents	360	1	35/60	210
	Children	360	1	53/60	318
Informed Consent II *(Parent-Attach. 16 or 18; Child-Attach. 26)*	Parent & Child	144	1	13/60	31
Session II (Retest)* *(Parent-Attach. 10,13; Child-Attach. 1-6)*	Parents	144	1	13/60	31
	Children	144	1	30/60	72
TOTAL		792			814

 * The parent and child will be tested simultaneously, either in the same room or in separate rooms near each other.

Annualized Cost to Respondents – Assessment of Executive Function for the NCS (LOI2-QUEX-6)

Data Collection Activity	Type of Respondent	Estimated Total Annual Burden Hours	Hourly Wage Rate*	Respondent Cost
Recruitment Screening Interview	Parents	66	$22.01	$1,453
Informed Consent I	Parent & Child	86	$22.01	$1,893
Session I (Test)	Parents	210	$22.01	$4,622
	Children	318	$22.01	$6,999
Informed Consent II	Parent & Child	31	$22.01	$682
Session II (Retest)	Parents	31	$22.01	$682
	Children	72	$22.01	$1,585
TOTAL		814		$17,916

* Based on the mean wages for all occupations. National Compensation Survey: Occupational wages in the United States May 2012, U.S. Department of Labor, Bureau of Labor Statistics.

☒ **Please check here after ensuring that all calculations have been verified**

NCS Formative Research Template for OIRA Clearance

Estimated Costs: Staff Hours: 1628 hours
                       Supervisor Hours: 407 hours

Attachments:
Attach 1: Touch Screen Tutorial
Attach 2: Dimensional Change Card Sort Test
Attach 3: Flanker Inhibitory Control and Attention Test
Attach 4: Picture Vocabulary Test
Attach 5: Picture Sequence Memory Test
Attach 6: Peg Tapping
Attach 7: WPPSI-IV Block Design
Attach 8: Woodcock-Johnson III Tests of Achievement
Attach 9: Family Demographic Questionnaire
Attach 10: CBQ-VSF+EF
Attach 11: Ages and Stages Social Emotional Questionnaire
Attach 12: Strengths and Difficulties Questionnaire
Attach 13: Q-ABC
Attach 14: Letter of Invitation
Attach 15: Recruitment Screening Interview
Attach 16: Consent Form Delve with 2nd session
Attach 17: Consent Form Delve with no 2nd session
Attach 18: Consent Form UofM with 2nd session
Attach 19: Consent Form UofM with no 2nd session
Attach 20: Univ of Minnesota IRB Original Application 01052011
Attach 21: Univ of Minnesota IRB Annual Renewal Approval 03062014
Attach 22: Univ of Minnesota IRB Change in Protocol Request 11062013
Attach 23: Univ of Minnesota IRB Change in Protocol Approval 11132013
Attach 24: Delve IRB Letter
Attach 25: Northwestern Univ IRB Letter
Attach 26: Child Assent Form
Attach 27: Univ of Minnesota IRB Change in Protocol Approval 06062014

☒ **Please check here after ensuring that the OMB #: 0925-0661 and Expiration Date: 06/30/2015 date have been inserted as first-page headers on each proposed instrument.**

☒ **Please check here after ensuring that the following OMB burden statement has been inserted as a first-page footer on each proposed instrument.**

Public reporting burden for this collection of information is estimated to average [insert estimated response time] minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. An agency may not conduct or sponsor, and a person is not required to respond to, a collection of information unless it displays a currently valid OMB control number. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to: NIH, Project Clearance Branch, 6705 Rockledge Drive, MSC 7974, Bethesda, MD 20892-7974, ATTN: PRA (0925-0661). Do not return the completed form to this address.