

**Pre-testing of Evaluation Surveys:
An Examination of the Intersection of Domestic
Human Trafficking with Child Welfare and Runaway
and Homeless Youth Programs**

**Information Collection Request
0970 - 0355**

Supporting Statement

Part B

October 2015

Submitted By:
Office of Planning, Research and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

7th Floor, West Aerospace Building
370 L'Enfant Promenade, SW
Washington, D.C. 20447

B1. Respondent Universe and Sampling Methods

Pretest Sample Size and Characteristics

We expect to pretest the screening tool with a convenience sample of 600 youths in multiple locations. The target population for the screening tool is youth involved in child welfare (CW) or runaway and homeless youth programs (RHY) ages 12-24. In order to find youth to participate, we are forming partnerships with foster care group homes, transitional living centers, RHY shelters, street outreach, and/or drop-in centers. Each site will identify clients who would be interested in participating in the study and then set up appointments for them to take the screening tool. At that time, once potential participants are read the assent/consent form, they can decide whether or not they want to participate. Thus, we will be collecting data from a convenience sample based on current and ongoing client/youth enrollment in the program/agency and interest in participating in the study. Because it is a convenience sample, the results from the tool will not be considered representative of youth in CW and RHY settings more generally, nor will they provide an accurate estimate of the prevalence of trafficking involvement in any particular setting or location, or among RHY and CW programs in general.

The target of 600 was set for the following reasons. The purpose of this exploratory project is to pretest a human trafficking screening instrument in multiple program settings and for varying populations of homeless, runaway and child welfare-involved youth. Multiple sites are critical for pretesting the instrument among a diverse set of youth in a diverse set of contexts. Specifically, the project team in consultation with HHS/ASPE has established that a minimum of three states with varying youth populations and two sites within each state—one serving homeless and runaway youth and one serving child welfare-involved youth—is necessary toward this end.

The instrument developed by the Urban Institute researchers consists, in its longest version, of approximately 20 items concerning youths' sexual and labor trafficking experiences. It has to date undergone pretesting for item content and administrative feasibility by the project team and fewer than 10 youth from the relevant population. Within each of the minimum six testing populations (i.e., RHY-youth and CW-youth in each of three states), adequate tests of the instrument's construct validity—established through exploratory and confirmatory factor analyses—depend on a minimum sample size of 100 and minimum subject-to-item ratio of 5, or 100 subjects for the 20-item instrument (Suhr, 2003; DeVellis, 2003; Tabachnick & Fidell, 2001). More generally, this minimum of 200 youth per state accords with that calculable from the simple sample size formula provided in Daniels (2013): with a desired confidence level of 95% ($Z=1.96$), estimated prevalence (P) of 15% trafficking (based on the recent Covenant House study), and expected error (d) of plus or minus 5% (see Daniels, 2013). Specifically, $n = Z\text{-squared} * P * (1 - P) / d\text{-squared} = 196$ youth per state. Collectively, this equates to an overall sample of 600 youth, or 200 youth from each state, as necessary to establish both the construct and external validity of the newly developed screening instrument.

We will work to ensure diversity in our sample, by CW vs. RHY involvement, age, gender, race/ethnicity, and sexual orientation. Below, we detail how we will work to ensure this

variation. In each case, the goal is not to conduct subsample analyses along any of these characteristics, but rather to ensure that the pretest is conducted across the full diversity of the target population.

To pretest the tool in both CW and RHY settings and among RHY and CW-involved populations, we will ensure that we obtain a large sample of youth from each type of program within each of the three locations. Specifically, we will work to ensure that we capture at least 75 RHY-involved youth and at least 75 youth from CW settings in each of our three study sites, out of the 200 total youth surveyed in each site. Thus, across all three locations, the number of RHY-involved youth we survey will range from a minimum of 225 (75x3) to a maximum of 375 (125x3), and the number of CW youth surveyed will have the same minimum (225) and maximum range (375).

To ensure that the tool is pretested across the broad age range we plan to target, we will work to ensure that—among the RHY youth surveyed in each state (min=75, max=125)—there are at least 15 youth from each of four age categories (12-14; 15-17; 18-20; 21-24).¹ We expect to encounter older youth primarily through RHY drop-in centers. Likewise, we will ensure that among the CW youth surveyed in each state (min=75, max=125)—there is a similar level of variation among the possible age categories within the age cap for child welfare involvement that applies in each site. We expect that this variation will be easy to obtain for youth in age category 15-17, so if we find that it is very difficult to find younger preteens and teens to survey, we will consider increasing the minimum age for our study during the data collection process. Likewise, if we are unable to identify enough youth over age 21 or 22, we may decrease the maximum age for our study.

As survey data are gathered over the course of the study, we will keep an eye on the gender, racial/ethnic, and sexual orientation breakdown of the sample, to ensure a reasonable distribution and representativeness of such youth. This monitoring will be done by observing the electronic data transmitted securely to Urban researchers as each new electronic tablet survey is completed and, for paper survey administration, through general discussions and inquiries with on-site program staff about the observed characteristics of youth who opt to take the survey (program staff would not review the survey instruments themselves, which would violate privacy, but rather would report on what they know about the characteristics—age, race/ethnicity, gender, and sexual orientation--of the youth they observe volunteering to take the self-administered or practitioner-administered surveys). Although we do not intend to require minimum numbers or percentages of any subgroup, if we see particularly low responses by any gender or racial/ethnic group or by lesbian, gay, bisexual, transgender, or questioning (LGBTQ) youth, we will take corrective action to target remaining survey recruitment efforts to reach those populations. We expect that our sites will naturally generate high numbers of responses across gender and racial/ethnic groups and include substantial shares of LGBTQ youth.

¹ We intend to examine the reliability and validity statistics for the group as a whole, and also to visually look to see how those statistics vary across subgroups defined by age, gender, etc. We do not anticipate conducting formal statistical tests to distinguish non-zero differences among subgroups, but we could do those if non-trivial differences emerge. Additionally, in the case that we are unable to recruit enough of a sample for each age group, we will re-assess our plan and proposed analyses.

Selecting the Pretest Sites

The pretest is to be conducted in several sites located in three states across the United States. In each of the three states, the *Life Experiences Survey* is to be administered to 200 runaway and homeless and/or child welfare-involved youth between the ages of 12 and 24, for a total sample of 600 youth. Currently, we are interested in administering the tool in Houston, TX; New York City, Westchester County, Nassau County, and Rochester, NY; and Milwaukee, WI. These locations offer substantial diversity in policy climate, geographic location, racial and ethnic composition of youth populations, and setting, including urban, suburban, and rural locations. Based on available publicly available information and informal individual conversations, we are currently building lists of RHY providers and CW group homes in each county with which to partner to administer the tool. The service environments in which pretesting will occur may include foster care group homes, transitional living centers, RHY shelters, street outreach, and/or drop-in centers.

Validation and Analysis

After completing data collection, we will take a number of steps to assess the internal validity and factorial reliability of the short and long versions of the tool pretested with youth in this convenience sample. First, we will examine issues such as the distributions of responses (e.g., to identify any problematic surveys where—for example—all answers are “yes,” or no response categories are ever chosen), whether any items are perfectly correlated meaning one could be dropped without loss of information, and whether survey timing is as anticipated and/or varied by youths’ age, demographic characteristics, or setting.

After these assessments, we will conduct analyses to assess the factorial validity, internal consistency reliability, sensitivity, and specificity of both the long- and short-form tools. First, we will assess the tools’ factorial validity to understand how the different question items correlate with one another, including whether they can be justifiably grouped into a single factor (or concept) measuring trafficking victimization overall and whether—for the long-form tool—certain subdomains also exist, such as fraud, force, coercion, or sexual trafficking victimization. Accordingly, we will use factor analysis to examine these interrelationships among question items. Factor analysis looks at the individual questions, or variables, and sees the variation among youths’ responses to those individual questions as a function of the underlying factors, or concepts, that item purportedly measures (plus error). In this way, factor analysis helps us tease out the most appropriate number of underlying factors to a set of question items. We will match individual questions to the appropriate factor(s) they appear to measure based on their “factor loadings,” or the amount of variance they appear to share with other questions measuring that same factor. Following recommendations in the field (Costello & Osborne, 2005) and based on our own expertise, questions with a factor loading of 0.4 or higher will be determined to effectively measure that particular factor, and all factors with three or more questions loading on them will be considered viable subdomains. From our review of factor analysis results, we will make an assessment as to the factorial validity of the long-form screener overall and with regard to any subdomains. For the short-form tool, we will assess the factor loadings of all questions on a hypothesized single factor measuring any trafficking experience, and we will examine the degree of correlation between individual items and any subdomains of the long-form screener.

Second, we will conduct a similar but statistically more simplistic assessment of the interrelationships among items included in the short- and long-form screeners by measuring the Cronbach's alpha of the factors and subdomains derived from the factor analyses. Cronbach's alpha measures the internal consistency, or reliability, of items included in a tool through an examination of their intercorrelations. Cronbach's alpha reliability coefficients with a value of 0.7 and higher will be determined to be acceptably high (Nunnally & Bernstein, 1994).

Third, we propose to conduct tests to give a preliminary sense of the degree to which trafficking-involved individuals are correctly identified as such (sensitivity, or the true positive rate) and the degree to which individuals who have no trafficking involvement are also correctly classified (specificity, or the true negative rate). This initial sense will inform future data collection to validate the tool under full OMB clearance. We will conduct these tests separately using the short- and long-form versions of the tool and the self-administered and practitioner-administered versions of the tool to compare their associated sensitivities and specificities. Once each screener's sensitivity and specificity has been calculated, these values will be combined into positive and negative likelihood ratios to assess the degree to which the tools correctly predict who has been trafficked and screen out those who have not.² These tests are technically carried out via simple binary classifications that compare individuals' determinations by a screener of interest as trafficking-involved or not against a "golden yardstick" assessment of their true trafficking involvement or not. Methods of validation include:

- *Practitioner Assessment.* For the 25% of the youth who receive a practitioner-administered tool, we will ask practitioners to write up their personal opinions as to a youth's trafficking involvement after they administer the tool, and we will use these opinions as an independent assessment of the youth's trafficking involvement.
- *Compare to Previously Validated Tool.* Wherever possible, we will compare youths' responses to the short- and long-form tools to a subset of the long-form questions taken directly from the previously validated Covenant House tool to serve as a measure of comparison reports.
- *Seeding the Sample.* If we select sites that already serve and know of youth who have been trafficked, we can incorporate some of these youth into our sample. A valid instrument should identify these youth as trafficking victims.
- *Cross-Validation with Another Trafficking Assessment.* If we select a site that is already simultaneously testing a diagnostic trafficking tool, which may be possible in New York and Houston, then we will attempt to coordinate administration of the *Life Experiences Survey* to youth who are also simultaneously being assessed by practitioners for trafficking victimization based on a separate diagnostic tool or observations. If it is possible to administer both tools to the same youth without compromising the goals of the present study, we will be able to compare the short- and long-form tools against the other assessments' results.
- *Prevalence of Known Correlates.* Few items adequately correlate with trafficking, meaning we can't collect a set of measures that will serve as proxies for direct questions

² Positive likelihood ratios are calculated as the sensitivity score divided by one minus specificity, and a value of 5 or higher indicates moderate to strong likelihood that the tool correctly identifies those who have been trafficked. Negative likelihood ratios are calculated as one minus the sensitivity score divided by specificity, and a value of 0.2 or less indicates moderate to strong likelihood that the tool correctly identifies those who have *not* been trafficked.

about trafficking. However, if trafficking victims have higher prevalence of certain characteristics, such as being LGBTQ, then the youth our tool identifies as trafficking victims should have higher rates of those characteristics than the youth our tool identifies as non-victims. We will identify such correlates wherever possible.

Lastly, using all of the information above, we will compare the responses to the short- and long-forms of the tool, to test the short form's adequacy as a substitute instrument. If modifications can be made to the language of the short-form questions to improve the scores based on information learned from the long-form item ratings, those modifications will be suggested for future piloting of the tool.

B2. Procedures for Collection of Information

At each pretest study site, we will administer the tool with young people between the ages of 12 and 24 who are currently receiving some type of service from the CW system or the RHY site. Surveys will be administered over a several-month period, and Urban staff will make trips of 4-5 days to each site to conduct survey administration training with practitioners and launch the first week of surveys.

During our trips to the testing sites, Urban staff will help recruit youth, train service provider staff, and oversee initiation of the pretest. After our visit, the local program staff will continue administering the tool to youth until a sample of 200 youth is obtained in each geographic site, across the various service providers and organizations with which we partner. We will ensure that this sample size is obtained within five months, recruiting additional testing locations if needed to increase the speed of data collection. If we deem it necessary, Urban staff will make a second site visit to assist with any further pretest administration or address any challenges that the staff have faced, or to launch the survey with additional service providers.

Data will be collected by administering a survey to youth through a computer or tablet, whenever possible. The data collection steps are as follows: First, the youth will complete the cognitive screener. If the individual passes the cognitive screener, the youth will then be asked if they are under the age of 18. Based on the administration method and the answer provided regarding age, the youth will see the appropriate consent form. After reading the consent form and consenting to continue the survey, the youth will take the rest of the *Life Experiences Survey*.

One-quarter of youth will take the tool as administered by a practitioner, who will read questions aloud, and record answers on the computer/tablet. Three-quarters of respondents will self-administer the survey using a computer/tablet. In either case, service providers will talk youth through the consent forms, and then youth will indicate whether they consent to continue with the study. If any locations are not open to using the electronic survey or do not have the necessary wifi capabilities to do so, we will also have a paper version available with clearly marked skip patterns. We will maintain a record of which locations used paper tools and which used electronic tools, and will note in our dataset of responses which came from paper versus electronic administration. Methods for securing data collected are detailed in Supporting Statement Part A.

We will also provide training and technical assistance throughout the pretesting study implementation process, including weekly check-in phone calls and/or e-mails to assess progress at each site and debrief about issues that may have arisen during survey administration. All practitioners involved in administering the tool will have multiple means (office phone, cell phone, e-mail) for contacting project staff as needed. In order to generate individualized familiarity and attention, we will designate one site liaison from our study for each site, at the research associate level, who will serve as the front line of contact with the site.

Prior to the start of the interview, subjects will be assessed for their appropriateness for eligibility and recruitment into the study. Before taking the survey, youth will take a cognitive screener. The cognitive screener is attached as Appendix B (Cognitive Screener). Those who are deemed ineligible based on the cognitive screener for reasons that would make the youth unfit for participating (e.g., they exhibit obvious signs of substance abuse or mental illness) will not be screened, and the practitioners will document in a special incident report all efforts that they make to provide appropriate referrals or assistance to that youth, following regular practices at the organization for referring youth in need to services. We do not anticipate that there will be a significant number of disqualifications at the cognitive screener level.

B3. Methods to Maximize Response Rates and Deal with Nonresponse

Expected Response Rates

We expect that some share of youth at each location will not be interested in participating in the study. Because our goal is to pretest our tool, we are not as concerned about achieving a representative sample than we would be if our aim were to describe the characteristics of the sample universe. Instead, our aim is to include a diverse set of youth who represent the range of characteristics of our target population, so that we can pretest the tool across a range of youth characteristics.

Maximizing Response Rates and Dealing with Nonresponse

To maximize response rates, as part of study recruitment youth will be informed of the importance of the study, that their responses will be kept private, and that they will receive a \$25 gift card in appreciation for their participation. If we find that response rates are lagging at any given site, we will work with that site to address recruitment challenges, review the procedures being used to recruit youth, and suggest revised methodologies where appropriate. As noted above, if we find that responses are lagging among any particular group, by gender, racial/ethnic group, or sexual orientation, we will similarly work with sites to try to boost participation among underrepresented groups.

B4. Tests of Procedures or Methods to be Undertaken

The purpose of this information collection request is to pretest our screening tool. The pretesting conducted under this clearance will result in an improved screening tool that can undergo future external validation under full OMB clearance.

B5. Individual Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

HHS has contracted with the Urban Institute to design the pretest tool instrument, organize the data collection, and analyze the data for validity and reliability. As noted above, some surveys will be collected by practitioners at the site. These individuals will be selected based on their knowledge and experience working with the target populations. The Principal Investigators of the study at the Urban Institute are:

Michael Pergamit, PhD
202-261-5276
mpergami@urban.org

Meredith Dank, PhD
347-404-7990
mdank@urban.org