

SUPPORTING STATEMENT
FOR OMB CLEARANCE
PART B

SUPPORTING HEALTHY MARRIAGE DEMONSTRATION EVALUATION
WAVE TWO DATA COLLECTION

ADMINISTRATION FOR CHILDREN AND FAMILIES
OFFICE OF PLANNING, RESEARCH AND EVALUATION

May 22, 2009

B. COLLECTION OF INFORMATION USING STATISTICAL METHODS

Respondent Universe and Sampling Methods

There is no change in the respondent universe or sampling methods for Wave 2. Respondents included in the second wave of information collection are all research sample member couples for the adult survey and a focal child within the household for the youth survey and in-home observation. The focal child in each family was identified previously.

The evaluation literature often discusses the appropriateness of the sample size for a study by focusing on the smallest program impacts that are likely to be detected with a specified level of confidence, assuming a sample of a given size and characteristics. These are usually called the program's "minimum detectable effects" (MDEs). Analysis of MDEs is also referred to as "power analysis," as it estimates the study's power to measure the effects it was designed to find.

Exhibit B1 shows the minimum detectable effects (MDEs) that can be achieved with different sample sizes using the survey data. The exhibit shows minimum detectable effect sizes (that is, the difference between the program and control group in terms of the number of standard deviations of the outcome) for a range of possible sample sizes for different types of analyses.

Exhibit B1

Minimum Detectable Effects for Various Sample Sizes in Supporting Healthy Marriage Follow-up Surveys

| <u>Size of Program and Control Group</u> | <u>Effect size</u> |
|--|--------------------|
| 160/160 | 0.28 |
| 320/320 | 0.2 |
| 640/640 | 0.14 |
| 960/960 | 0.11 |
| 2560/2560 | 0.17 |

NOTE: MDEs are for two-tailed tests at 0.1 significance with 80 percent power.

The exhibit shows MDEs for several sample sizes: (1) 160 in each research group, which could represent a subgroup of half of the sample in a site, (2) 320 in each research group, representing a single site, (3) 640 in each research group, representing results for the two curricula that will be used in two sites, (4) 960 in each research group, representing results for the curriculum that will be used in three sites, and (5) 2,560 in each research group, representing results for all sites pooled.

As the exhibit indicates, the MDE for a sample of 320, representing the sample in a single site, is 0.20 standard deviations. This means that if the true effect of an intervention is 0.20 standard deviations, then the difference in survey-based outcomes between program and control groups would be statistically significant in 80 percent of experimental tests of that intervention. Compared with many marital interventions studied using random assignment with middle-class white couples, a short-term impact of 0.20 standard deviations is not especially large. Meta-analyses of marriage education and marital and family therapy have found average effect sizes at post-program assessments of 0.50 standard deviations or more.

If sites are pooled, the study has a much better chance of finding statistically significant impacts on survey-based outcomes. For the two curricula being tested in two sites, for example, the MDE is about 30 percent lower when the two sites are combined than when they are looked at separately. For the curriculum being tested in three sites, the MDE is more than 40 percent smaller when the three sites are pooled. Finally, since it might be difficult to find statistically significant impacts in any individual site, we plan to estimate results pooling data from all eight sites. This will reduce the MDE by nearly two thirds.

MDES FOR DIRECT CHILD ASSESSMENTS AND YOUTH-REPORTED QUESTIONNAIRE

The direct child assessments and youth survey will be fielded with approximately 800 children in each site, divided equally between the program and control groups. Our goal is to achieve a 72 percent response rate, resulting in complete assessments and youth surveys. Children between the ages of two years and eight years, five months old at the 30-month follow-up will receive direct child assessments. Children between the ages of eight years, 6 months and 17 years old at the 30-month follow-up will receive a youth survey. Based on the sample intake characteristics of children of couples randomly assigned in the study thus far and anticipating a 72 percent response rate, we expect that approximately 3,225 children will complete the direct child assessments and 1,383 children will complete the youth survey across the 8 sites for wave 2, at 30 months after random assignment.

Exhibit B2 shows the minimum detectable effects (MDEs) that can be achieved with different sample sizes using the youth survey in the wave 2 follow-up.

Exhibit B2

Minimum Detectable Effects for Various Sample Sizes in Supporting Healthy Marriage 30-month Follow-up

| Youth Survey | |
|--|--------------------|
| Size of Program and Control Group | Effect Size |
| 317/317 | 0.20 |
| 475/475 | 0.16 |
| 950/950 | 0.11 |

NOTE: MDEs are for two-tailed tests at 0.10 significance with 80 percent power.

Exhibit B2 shows MDEs for several sample sizes: (1) 950 in each research group, which could represent the sample size for the results of youth survey outcomes for all sites pooled; (2) 475 in each research group, representing results for a subgroup of half of the sample size for youth survey outcomes for all sites pooled; and (3) 317 in each research group, which could represent the results for a subgroup of one third of the sample size for the youth survey outcomes.

As the exhibit indicates, the MDE for outcomes using the youth survey with data that is pooled across all 8 sites (the smallest sample size in the exhibit) is 0.11 standard deviations. This means that if the true effect of an intervention is 0.11 standard deviations, then the difference in youth-reported outcomes between program and control groups would be statistically significant in 80 percent of experimental tests of that intervention. Compared with interventions that directly target child well-being or family functioning, an impact on child outcomes of 0.11 standard deviations is a small effect size and somewhat smaller than the effect sizes found in recent evidence from a small-scale experimental evaluation of a relationship skills and parenting program (e.g. Cowan, et al., in press), suggesting that the current study will have adequate power to detect intervention effects on the pooled and subgroup samples.

The other child-reported data to be collected during the second wave of follow-up will be the direct child assessments, which will be administered to about two thirds of the focal children in the full sample. Exhibit B3 shows the minimum detectable effects (MDEs) that can be achieved with different sample sizes using the direct child assessments at the 30-month wave 2 follow-up.

Exhibit B3

Minimum Detectable Effects for Various Sample Sizes in Supporting Healthy Marriage 30-month Follow-up

| Direct Child Assessments | |
|--|--------------------|
| Size of Program and Control Group | Effect Size |

| | |
|-----------|------|
| 179/179 | 0.26 |
| 269/269 | 0.21 |
| 359/359 | 0.19 |
| 538/538 | 0.15 |
| 1075/1075 | 0.11 |
| 1613/1613 | 0.09 |

NOTE: MDEs are for two-tailed tests at 0.10 significance with 80 percent power.

As is the case with the youth survey, because the age-specific samples for the direct child assessment are somewhat smaller than the sample for the adult surveys, we propose examining impacts on child-level outcomes using data that is pooled across the sites. For the most part, we propose examining impacts on child-level outcomes separately by the direct child assessment technique/task. As such, the impacts on child outcomes measured with direct child assessments would be analyzed separately for children who are between the ages of 2 to 3 years, 5 months old from children who are between the ages of 3 years, 6 months to 8 years, 5 months old. We also will explore whether it is possible to examine impacts on children's self-regulation outcomes by pooling data across assessment techniques that are used with children of different ages to maximize our sample sizes and power to detect intervention effects.

The exhibit shows MDEs for several sample sizes: (1) 1,613 in each research group, which could represent the sample size for the results for direct child assessment outcomes for all sites pooled, across both age ranges (this would be the sample if we are able to pool data on self-regulation that is gathered from different assessment techniques across children of different ages); (2) 1,075 in each research group, representing results for all children 3 years, 6 months to 8 years, 5 months old for all sites pooled; (3) 538 in each research group, which represents the sample size of children ages 2 to 3 years, 5 months old across the sites, or half of the subgroup of children ages 3 years, 6 months to 8 years, 5 months old across the sites; (4) 359 children in each research group, representing one third of the pooled sample of children ages 3 years, 6 months to 8 years, 5 months old; (5) 269 in each research group, representing half of the pooled subgroup of children ages 2 to 3 years, 5 months old; and (6) 179 in each research group, or the equivalent of about one third of the pooled group of children ages 2 to 3 years, 5 months old.

As the exhibit indicates, the MDE for outcomes using the direct child assessments are from 0.09 to 0.26 standard deviations. That is, for example, if the true effect of an intervention is 0.09 standard deviations, then the difference in child assessment outcomes between program and control groups would be statistically significant in 80 percent of experimental tests of that intervention. Compared with interventions that directly target child well-being or family functioning, MDEs in this range are small to modest in size and in line with recent evidence from a small-scale experimental evaluation of another relationship skills and parenting program (e.g. Cowan, et al., in press).

B2. Procedures for Collection of Information

The procedures for administering the survey in wave 2 are the same as in wave 1 (12 months post random assignment). The activities to be undertaken during the in-home activities are different but the procedures are similar. The approach is summarized below.

- About 30 months following their random assignment, couples enrolled in the SHM study will be sent a letter reminding them of their participation in the SHM study and informing them that they will soon receive a phone call from a representative of the survey research firm who will want to interview them over the phone about their marriage and children and conduct direct child assessments or a youth survey with one of their children.
- For wave 2 of the adult survey, interviewers will call the specified contact numbers for SHM study participants and administer the 50-minute follow-up survey to all willing participants approximately 30 months after study participants first entered the study.
- Upon completion of the questionnaire, the interviewer will ask parents with focal children who are eight years, 6 months old and older at the 30-month follow-up if they would allow their child to complete a 30-minute questionnaire. For children 11 years old and older, the interviewer will ask the parent for contact information for the focal child to complete the survey over the telephone. For children between the ages of eight years, 6 months

and 11 years old, the interviewer will ask the parent if there is a suitable time when the child will be available so that an interviewer can complete the youth survey in person. If the parent agrees to allow his/her child to participate in the data collection effort, a survey firm representative will make arrangements to go to the family's home to complete the youth interview in person or over the telephone, depending upon the age of the child at the time of the 30-month second wave.

- Upon completion of the 30-month survey, the interviewer will ask parents with focal children who are currently eight years, five months old and younger if they would allow their child to complete 30 minutes of direct child assessments. If the parent agrees to allow their child to participate in the data collection effort, a representative will make arrangements to go to the family's home to complete the direct child assessments in person.

Interviewer training. Training will take place close to the time when the first cohort of research subjects reaches the 30-month anniversary of their random assignment date. The same field interviewers who administer wave 2 of the survey will receive extensive training to conduct the direct child assessments and youth surveys. In the past, this has typically involved training sessions which last about five to seven days in total. The training will include an item-by-item or task-by-task review of the survey instrument, practice interviews and administrations, and critiques of those interviews. Finally, each interviewer will undergo a certification process prior to fielding to ensure that the interviewer is qualified to set up the survey, direct child assessments and youth survey. The interviewers will also be monitored during early interviews and will be subject to periodic reviews over the course of fielding the data collection instruments to ensure that they are following procedures and protocols with a high degree of fidelity.

In addition, the interviewer training will include training on a protocol for handling adverse events while fielding the data collection instruments. The protocol was developed for the 12-month follow-up, and has already been approved by relevant Internal Review Boards.

All interviewers will sign a confidentiality pledge during training. They will be instructed on the protocols of maintaining confidentiality and told that breaches of confidentiality will lead to dismissal.

B3. Maximizing Response Rates

The goal will be to administer the wave 2 survey to all sample members in each site, and the direct child assessments and youth questionnaire to all of the focal children, as defined by their age. Procedures for obtaining the maximum degree of cooperation include:

- Conveying the purposes of the study and follow-up data collection efforts to respondents so they will thoroughly understand the purposes of the data collection and perceive that cooperating is worthwhile;
- Providing a toll-free number for respondents to use to ask questions about the data collection efforts and procedures;
- Training site staff to be encouraging and supportive, and to provide assistance to respondents as needed;
- Training interviewers to maintain any pre-existing one-on-one personal rapport with respondents
- Offering appropriate payments to respondents; and
- Self-administered CATI/CAPI procedures to maximize respondents' privacy during the administration of sensitive questions.

Collectively, these methods have also been shown to positively affect response rates by enabling people with limited literacy skills (particularly important given that the study sample disproportionately speaks English as a second language) to respond to sensitive questions while maintaining their privacy (Belcher, et al., 2001).

B4. Pre-testing

The 12-month survey was developed with extensive pretesting and cognitive testing. The wave 2 changes/survey will undergo a round of pretesting to provide information about the length of the instrument, with the goal of improving the

quality of the data that the instrument would yield. Because many of the items were directly drawn from the 12-month survey, we expect few if any revisions to the instrument will be needed as a result of this pretesting.

The protocols for the direct child assessments are drawn from prior research and protocols that have been used to administer these kinds of tasks in other large-scale program evaluation projects (e.g., the Enhanced Services for the Hard-to-Employ (HtE) project [OMB No. 233-01-0012]; the Head Start CARES project [OMB No.0970-0364]; the Building Strong Families project [OMB No. 0970-0344]). We expect to pretest these protocols with a small number of children of various ages to provide information about the length of the various assessments and streamline the protocols as needed.

Items for the youth survey were drawn from prior large-scale experimental and non-experimental research. We expect to pretest this questionnaire with a small number of appropriately-aged children to provide information about the length of the instrument and to identify areas in which we might streamline the questionnaire if needed.

B5. Consultants on Statistical Aspects of the Design

There are no outside consultants on the statistical aspects of the design. We have drawn on the considerable statistical and methodological expertise of the SHM team members including Charles Michalopoulos and Howard Bloom of MDRC and Larry Orr and David Fein of Abt Associates.

References

- Belcher, H.M., Butz, AM, Pulsifer M., Marano, N. 2001. Effectiveness of a home intervention for perceived child behavioral problems and parenting stress in children with in utero drug exposure. *Archives of Pediatric and Adolescent Medicine*. Vol.155:1029-1037.
- Cowan, P.A., Cowan, C.P., Pruett, M.K., and Pruett, K. In press. Promoting fathers' engagement with children: Preventive interventions for low-income families. *Journal for Marriage and Family*.