Table C.1 in the May 2007 Sample Design Recommendation Memo (***Volume II***) provides an allocation scheme that was used in our analysis foe the NSCAW II Consultant Group to illustrate the effect of oversampling on precision. The estimated design effects associated with this scheme are provided in Table C.2. The overall design effect is estimated at 4.6, although this may be somewhat of an underestimate. Changing the allocation to domains can also change the design effect. Nonresponse, frame undercoverage and other weighting adjustments can inflate this number. In addition, it will not be possible to achieve exactly the sample sizes specified by the optimization algorithm. Much depends on the accuracy of projections of the domain sizes in each PSU. Work is currently underway to evaluate the proposed NSCAW II design in light of our expectations and realizations for NSCAW I, in order to provide better projections of the NSCAW II design effects.

Statistical Power. It is impossible to specify the statistical power of NSCAW analysis overall because it will include so many different research questions, variables and subpopulations. The proposed cohort size of 5,700 child interviews is slightly higher than the NSCAW I survey, which had 5,501 completed key respondent interviews. Based on experience with the NSCAW I analysis, this sample was adequate for many types of analysis that were conducted, both for cross-sectional and longitudinal analysis. As an example, for the descriptive estimates contained in Wave 1 Analysis Report (DHHS, 2005), the coefficients of variation (cv's) of the estimates never exceeded 20%. Further, for comparing estimates between independent subgroups or for comparing estimates of the same subgroup at two points in time, the precision of the comparisons was quite adequate for the eight planned domains of analysis as well as for a number of ad hoc domains.

To illustrate the statistical power of the NSCAW II sample design (Note: The allocation changed after the Consultant Group meeting in May 2007), ***Exhibit B.1-2*** shows sample sizes for the five sampling domains, their corresponding effective sample sizes, and the power for detecting small effect sizes. These allocations are approximations, and may change based on the achieved sample. The effective sample sizes in the exhibit were computed as the actual sample size divided by the appropriate design effect. The design effect for the total sample is 5.6[1]. It varied by domain from 2.3 to 2.9. These design effects, which account for both clustering and the unequal weighting roughly on par with those observed for NSCAW I. Power is based on a two-sided test of size $\alpha = 0.05$ comparing two independent means or proportions; for example, the mean total score on the CBCL for children in two different domains. These calculations assume that a "small effect size" (i.e., an effect size of 0.2 on Cohen's effect size scale; Cohen, 1988) is to be detected.

[1] The design effect of 5.6 is higher than the 4.6 design effect projected for the May 2007 Consultant Group Meeting. This is due to changing the allocation scheme used in the Table C.1 in the Sample Design Recommendation Memo (***Volume II***) to the one in Exhibit B.1-2. The allocation in this exhibit is the final allocation scheme that will be used in the study.

Exhibit B.1-2 Sample Size, Effective Sample Size and Power by Domain

NSCAW II Sampling Domain				Actual N	Effective N	Power for Small Effect Size
Age	Open or Closed		In Foster Care (FC)?			
Less than 1 year	Open		Yes	1,100	470	0.86
1-17.5 years	Open		Yes	1,100	432	0.84
All Ages	Closed		No	1,300	522	0.90
Less than 1 Year	Open		No	1,100	429	0.83
1-17.5 years	Open		No	1,100	384	0.79
All Domains				5,700	1,020	0.99

The power shown in ***Exhibit B.1-2*** represents a worst case scenario since we have assumed independent means and small effect sizes. For example, increasing the effect size from 0.2 to 0.35 (which is still in the small range according to Cohen's rule) increases the smallest power in the table from 0.79 to 0.99. For dependent means (e.g., change estimates) the power may be closed to 1. For example, assuming a moderate correlation between the means will increase smallest power in the table to from 0.79 to 0.99. For medium and larger effect sizes (i.e., effect sizes of 0.50 and 0.80, respectively) the power is close to 1 for testing for differences between independent means. This analysis suggests that even for small subgroups in the population, the design provides adequate power for a wide range of analyses.