

Request for OMB Clearance
Evaluating Student Need for Developmental or Remedial Courses
at Postsecondary Education Institutions

(Formerly titled Survey of Placement Tests and Cut-Scores in Higher Education Institutions).

Supporting Statement for Paperwork Reduction Act Submissions

REVISED 3-30-11

Supporting Statement B: Statistical Methodology for Probe study

B.1. Respondent Universe for Probe study

The sampling frame or respondent universe from which the sample of institutions for the probe study will be drawn will be constructed from the most recent Institutional Characteristics (IC) component of the Integrated Postsecondary Education data System (IPEDS) maintained by the National Center for Education Statistics (NCES)¹. The 2009-10 IC file is currently available and will be used to select the sample for this study. However, at the time of the original OMB submission, the 2008-09 IC was available and it was used for the description of the sampling frame and sample selection for the “pilot” (i.e., feasibility study) and full-scale studies (note that the distributions for the 2008-09 IC file and the 2009-10 IC file are very similar).

The 2008-09 IC file contains over 7,100 Title IV postsecondary institutions. To be eligible for the study, 2-year and 4-year degree granting institutions must be located in the United States, and must offer a bachelors’ program or lower, although the institution may also offer graduate degrees. Thus, institutions offering only post-graduate or first-professional degrees will be excluded from the sampling frame. Administrative units are not eligible for the study because these entities do not provide instruction to students. Furthermore, like institutions in outlying territories, non-Title IV institutions are excluded because they are different from Title IV institutions. Non-Title IV institutions are not eligible for Pell grants and they generally yield very low survey response rates as a group. Thus, when describing the universe of postsecondary institutions, the NCES convention is to report separately for Title IV and non-Title IV institutions. As with many other sample surveys of postsecondary institutions, the current study excludes non-Title IV institutions because these institutions are generally too few in the sample to be reported separately.

Table B-1 summarizes how the frame of institutions will be constructed. As indicated in the last row of the table, 4,148 institutions meeting the eligibility criteria for the study will be included in the sampling frame, of which 2,458 will be 4-year institutions and 1,690 will be 2-year institutions. Table B-2 summarizes the distribution of eligible institutions in the sampling frame by level and control.

¹ At the time the original OMB submission, the 2008-09 IPEDS was the most current edition available. The 2009-10 IPEDS has since been released and will be used for sample selection. Thus, while the general approach to sampling will remain the same as described herein, the actual population and sample counts may differ slightly from those presented in the tables in this document.

Table B-1. Summary of frame construction for probe study

Description of institutions in 2008-09 IPEDS IC File	No. of			
	institutions	4-year	2-year	Other
All institutions	7,126	2,906	2,299	1,921
Institutions in United States (excluding outlying territories)	6,961	2,837	2,274	1,850
Institutions in United States offering undergraduate classes	6,649	2,541	2,274	1,834
Non-degree granting	2,377	6	537	1,834
Degree granting	4,272	2,535	1,737	n/a
Non Title IV	45	32	13	n/a
Title IV*	4,227	2,503	1,724	n/a
Administrative unit†	79	45	34	n/a
Non-administrative unit **	4,148	2,458	1,690	n/a

* Includes 4 US Service academies.

† Pending review, all or most of these will be deleted from the sampling frame.

** Count includes 52 cases with missing enrollment data. Unless deemed ineligible, all will be included in the final sampling frame.

Table B-2. Distribution of eligible postsecondary institutions to be included in the sampling frame, by level and control

Level	Control	Number of institutions with nonmissing enrollment in sampling frame **	Number of institutions in sampling frame
4-year	Public		
	Private, not-for-profit		
	Private, for-profit		
2-year	Public		
	Private, not-for-profit		
	Private, for-profit		
Total institutions (with non missing enrollment)*			

* The counts given in this table are based on tabulations from the 2008-09 IPEDS Institutional Characteristics (IC) file. The 2009-10 edition of the IPEDS is currently available and will be used for sampling; thus, the final population counts will differ slightly from those shown.

** The counts in this column exclude 52 eligible institutions in the 2008-09 IC file that have missing values for enrollment. However, such institutions will be included in the final sampling frame and given appropriate chances of selection. The resulting population counts including these 52 cases are shown in the last column of this table.

B.2. Statistical Methodology

Sample Design and Selection for the Feasibility study

The feasibility study was conducted in fall 2010. Since a primary purpose of the feasibility study was to identify potential data collection problems during the probe study, the goal of the sample design was to ensure that a diverse cross-section of postsecondary institutions in the United States were included in the sample. Thus, in addition to institutional control and level, the sample was stratified by characteristics such as size of institution and highest level of degree offering. Within a sampling stratum, institutions were selected at rates that are approximately proportional to the square root of their enrollment. The sample size of 120 institutions for the feasibility study was designed to yield about 100 respondents. Although a measure of selectivity was not used as a stratifier in sample selection, the sample of 4-year institutions included in the feasibility study covered a broad range of institutions with respect to selectivity.

Table B-3 summarizes the sample sizes for the feasibility study and the corresponding numbers of respondents by major subgroups defined by institutional level and control. As can be seen in the table, the response rate for all institutions was 86 percent, and exceeded 90 percent for public and private not-for-profit institutions. Response rates for private for-profit institutions were much lower, especially for 2-year private for-profit institutions which yielded the lowest response rate among all sectors. While these results suggest that response rates for the private for-profit institutions may be lower compared with public and private not-for-profit institutions, the small sample sizes and concomitant sampling error, make it difficult to extrapolate an exact value to the probe study and should be interpreted with caution. Thus, while the results in table B-3 can be used to provide general guidance, they cannot be used literally to estimate the sample yields for the probe study.

Table B-3. Sample sizes, number of respondents, and response rates for the feasibility study, by control and level

Subgroup	Number sampled	Ineligible*	Feasibility study Respondents	Response rate (%)**
<i>Total sample</i>	120	8	96	86
<i>Public</i>	60	0	54	90
4-year	30	0	27	90
2-year	30	0	27	90
<i>Private, not-for-profit</i>	40	2	35	92
4-year	30	1	27	93
2-year	10	1	8	89
<i>Private, for-profit</i>	20	6	7	50
4-year	10	2	5	63
2-year	10	4	2	33

*Closed institutions and those that did not offer programs for undergraduate students.

** Unweighted response rate among eligible institutions..

Sample Design and Selection for the Probe study

Sample Allocation for the Probe study

One of the goals of the sample design is to ensure that the sample sizes are sufficiently large to enable the detection of significant differences of 0.2 standard deviations in assessment score scales between major subgroups to the extent feasible. Table B-4 summarizes the minimum detectable effect size (i.e., the difference expressed as a multiple of the standard deviation) between subgroups of various sizes and a range of design effects. The design effects will arise primarily because of the varying probabilities of selection with which the institutions will be drawn into the sample. Depending on the magnitude of the design effect, it can be seen that each subgroup would have to include between 400-500 respondents to be able to detect an effect size of 0.20 (i.e., 0.20 standard deviations). Larger effect sizes of 0.40 can be detected with subgroups sample sizes of 100-200. Only very large effect sizes of 0.60 or greater can be detected with subgroup sample sizes as small as 50 to 60.

Table B-4. Minimum detectable effect sizes* between subgroups for selected sample sizes and design effects

Subgroup sample size [†]	Design effect		
	1.10	1.20	1.30
800	0.15	0.15	0.16
700	0.16	0.16	0.17
600	0.17	0.18	0.18
500	0.19	0.19	0.20
400	0.21	0.22	0.23
300	0.24	0.25	0.26
200	0.29	0.31	0.32
150	0.34	0.35	0.37
100	0.42	0.43	0.45
60	0.54	0.56	0.58

* Entries are detectable effect sizes for a two-sided test with alpha = 0.05 and power = 0.80.

† Sample size per subgroup.

In addition to the specified precision goals, disaggregation of the sample into subgroups defined by control and level will also be important for analysis. However, this will not be possible for very small subgroups without unduly reducing the overall efficiency of the sample. For example, it can be seen in Table B-2 that most of the private 2-year institutions are for-profit institutions. Since the private 2-year institutions tend to be small (fewer than 500 students), it does not seem useful to separate the 2-year private for-profit and not-for-profit institutions from either a statistical or policy perspective. Among the private 4-year institutions, on the other hand, there are several extremely large for-profit institutions (e.g., DeVry, University of Phoenix) that are inherently different from the traditional private not-for-profit institutions. In this case, separating for-profit 4-year institutions from the others would be desirable to provide meaningful comparisons.

In view of the above considerations, we propose to select the numbers of institutions indicated in the second column of Table B-5. Assuming an overall response rate of 85 percent, the proposed sample sizes will yield the numbers of respondents shown in the column of the table headed “Expected Respondents.” The assumed 85 percent response rate is broadly in line with the overall results from the feasibility study (see Table B-3). As indicated previously, we are cautious in our interpretation of data in making projections from the feasibility study, given the limitations inherent in its objectives and design. Data

from the feasibility study suggest that there might be differences in response rates among particular subgroups of institutions. Specifically, response rates for 2-year and 4-year for-profit private institutions might be lower, leading to possible shortfalls in the expected number of respondents for these subgroups. Since these types of institutions are included in the sample to ensure that they are appropriately represented in national estimates rather than for the purpose of separate subgroup analyses, it does not seem worthwhile to assume a lower response rate for these subgroups for sampling purposes.

Table B-5. Proposed target sample sizes for the probe study, by control and level

Subgroup	Number sampled	Expected Respondents*
Total sample	1,668	1,418
<i>Public</i>	936	796
4-year	468	398
2-year	468	398
<i>Private, not-for-profit</i>	490	417
4-year	468	398
2-year	22	19
<i>Private, for-profit</i>	242	206
4-year	160	136
2-year	82	70

* Assumes 85 percent response rate.

Note: Details may not sum to total due to rounding.

Despite some uncertainty about the response rates that can be achieved in the probe study, it nonetheless appears that it will be possible to detect an effect size of 0.20 for comparisons between public vs. private not-for-profit institutions, between public 4-year and public 2-year institutions, and between public 4-year and private not-for-profit 4-year institutions (Table B-4). Effect sizes of 0.40 or larger can be detected between smaller subgroups; e.g., comparisons between private not-for-profit and private for-profit institutions or between private 4-year not-for-profit institutions and private 4-year for-profit institutions.

The thresholds used for NCES’s reports will be applied to suppress survey data for the probe study; i.e., cells with less than 3 cases in the numerator and 30 cases in the denominator will be suppressed. In addition, estimates will be suppressed if the coefficient of variation is equal to or greater than 50 percent. Estimates with a coefficient of variation that is equal to or greater than 30 percent but less than 50 percent will be flagged to alert readers that the data should be interpreted with caution. As shown in Table B-5, the sample size of 22 for the probe study will be inadequate to compare private not-for-profit 2 year institutions (note that there are only 92 institutions in the frame). In addition, it may not be possible to compare data for private for-profit 2-year institutions and private for-profit 4-year institutions in the probe study.

Sample Selection Procedures for the Probe study

To permit the efficient allocation of the sample to meet the goals stated above, the sampling frame will be stratified by size of institution, level, and type-of-control, along with other institutional characteristics such as highest level of degree offering. Implicit stratification within major strata by geography and minority status of institution will also be employed to ensure that all regions and types of institutions are appropriately represented in the sample. Within a stratum we will sample institutions at rates that are

approximately proportional to the square root of their enrollment. Table B-6 shows the distribution of the proposed sample by each of the 39 sampling strata using this method. A similar design has been used to select institutions for the Postsecondary Quick Information System (PEQIS) and is expected to be efficient for the proposed survey on the use of placement tests and cut scores. Note that institutions with enrollments of 10,000 or more are selected with certainty under this design.

Table B-6. Distribution of postsecondary institutions to be included in the sampling frame

Stratum	Level	Type of Control	Highest level of offering	Enrollment size class	Number of institutions in frame*	Number of institutions in sample
1	4-year	Public	Doctorate	<3,000	20	7
2				3,000 to 9,999	63	42
3				10,000 to 19,999	94	94
4				20,000+	106	106
5			Masters	<3,000	57	19
6				3,000 to 4,999	51	34
7				5,000 to 9,999	87	65
8				10,000+	42	42
9			Bachelors	<1,000	19	5
10				1,000 to 2,999	45	15
11				3,000 to 9,999	38	25
12				10,000+	14	14
13	4-year	Private (not-for-profit)	Doctorate	<500	47	5
14				500 to 2,999	100	33
15				3,000 to 9,999	152	91
16				10,000+	53	53
17			Masters	<500	81	9
18				500 to 999	92	31
19				1,000 to 2,999	269	90
20				3,000 to 9,999	92	55
21			Bachelors	10,000+	8	8
22				<500	151	17
23				500 to 999	107	21
24				1,000 to 2,999	146	49
25	3,000+	13	7			
26	4-year	Private (for-profit)	—	<100	28	6
27			100 to 499	163	41	
28			500+	281	112	
29	2-year	Public	—	<1,000	113	16
30			1,000 to 2,999	288	72	
31			3,000 to 4,999	201	80	
32			5,000 to 9,999	241	121	
33			10,000+	179	179	

See notes at end of table.

Table B-6. Distribution of postsecondary institutions to be included in the sampling frame (continued)

Stratum	Level	Type of Control	Highest level of offering	Enrollment size class	Number of institutions in frame*	Number of institutions in sample
34	2 year	Private (nonprofit)	—	<500	67	13
35				500 to 999	20	7
36				1,000+	5	3
37	2 year	Private (for profit)	—	<500	369	41
38				500 to 999	154	31
39				1,000+	40	10
Total					4,096	1,668

* The population counts given in this table are based on initial tabulations from the 2008-09 IPEDS Institutional Characteristics (IC) file. Fifty-two eligible institutions in the 2008-09 IC file have missing values for enrollment and are not included in this table. Note that the 2009-10 IPEDS is now available and will be used for sample selection. Thus, the final population and sample counts will differ from those shown.

Note: Details may not sum to total due to rounding.

For the 4-year institutions, measures of selectivity will be used, as appropriate for the interpretation of the results from the survey. Thus, it is important that the sample include institutions covering a broad range of selectiveness. Applying the algorithm developed by Cunningham (2005),² we have estimated the numbers of institutions to be included in the sample by various categories of selectivity. The numbers of institutions in the frame and the corresponding expected sample sizes are shown in Table B-7. As can be seen in this table, a wide range of institutions with respect to selectivity is expected to be included in the sample. Note that other methods of defining selectivity can also be employed. The results in table B-7 are intended to provide rough orders of magnitude of the sample sizes to be expected under the proposed sample design.

²Cunningham, A.F. (2005). *Changes in Patterns of Prices and Financial Aid* (NCES 2006-153). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Table B-7. Number of institutions in the frame and expected sample sizes, by selectivity among 4-year institutions

Selectivity (based on Cunningham method)	Level	Number of institutions in frame	Number of institutions in sample
Most/Very Selective	Public	106	93
	Private	249	129
Moderately Selective	Public	316	248
	Private	560	213
Minimally Selective	Public	96	58
	Private	523	155
Open Admissions	Public	118	69
	Private	451	131
Total		2,419	1,095

Note: Details may not sum to total due to rounding.

Expected Levels of Precision for the Probe study

Table B-8 summarizes the approximate sample sizes and standard errors to be expected under the proposed design for selected subgroups. Since the sample sizes in Table B-8 are based on preliminary tabulations of the 2008-09 IC file, the actual sample sizes to be achieved may differ from those shown. Also, it is important to note that the sample sizes represent the expected numbers of completed questionnaires with eligible institutions, and not the initial numbers of institutions to be sampled. The standard errors in Table B-8 reflect an overall design effect of 1.3. The design effect arises primarily from the use of variable sampling fractions across the major sampling strata and differential nonresponse weighting adjustments (see description under **Estimation and Calculation of Sampling Errors**). In particular, the design effect reflects the fact that under the proposed stratified design, large institutions will be sampled at relatively higher rates (i.e., have smaller sampling weights) than small institutions. In fact, some very large institutions will be selected with certainty or near certainty. Subgroups consisting solely of such institutions will have negligible standard errors because the finite population correction (fpc) for these subgroups will be 0 or close to 0. Hence, the entries in the table should be viewed as rough upper bounds on the standard errors to be expected from the survey.

The standard errors in Table B-8 can be converted to 95 percent confidence bounds by multiplying the entries by 2. For example, an estimated proportion of the order of 20 percent ($P = 0.20$) for public 4-year institutions will be subject to a margin of error of ± 4.6 percent at the 95 percent confidence level. Similarly, an estimated proportion of the order of 50 percent ($P = 0.50$) for 4-year institutions requiring test scores for admission will be subject to a margin of error of ± 4.4 percent at the 95 percent confidence level.

Table B-8. Expected standard error of an estimated proportion under proposed design for selected analytic domains, by control and level

Domain (subset)	Expected sample size*	Standard error† of an estimated proportion equal to ...		
		P = 0.20	P = .33	P = .50
Total	1,418	0.012	0.014	0.015
<i>Public</i>	796	<i>0.016</i>	<i>0.019</i>	<i>0.020</i>
4-year	398	0.023	0.027	0.029
2-year	398	0.023	0.027	0.029
<i>Private, 4-year</i>	534	<i>0.020</i>	<i>0.023</i>	<i>0.025</i>
Not-for-profit	399	0.023	0.027	0.029
For-profit	136	0.039	0.046	0.049
<i>4-year schools</i>	931	<i>0.015</i>	<i>0.018</i>	<i>0.019</i>
Requires test scores	620	0.018	0.022	0.023
Has open admissions	144	0.038	0.045	0.048
Most/Very Selective**	188	0.023	0.027	0.028
Moderately Selective	392	0.023	0.027	0.029
Minimally Selective	181	0.034	0.040	0.042
Open Admissions	170	0.035	0.041	0.044

* Expected number of responding eligible institutions, assuming response rate of 85 percent. The standard errors given in this table are given for illustration. Actual standard errors may differ from those shown.

† Assumes unequal weighting design effect of 1.3. For subgroups consisting of institutions selected with certainty, the standard errors will be smaller than those shown.

** Standard errors include an approximate finite population correction to reflect the fact these institutions will be selected at relatively high rates.

Estimation and Calculation of Sampling Errors for the Probe study

For estimation purposes, sampling weights reflecting the overall probabilities of selection and adjustments for nonresponse will be attached to each data record. To properly reflect the complex features of the sample design, standard errors of the survey-based estimates will be calculated using jackknife replication. Under the jackknife replication approach, 50 subsamples or "replicates" will be formed in a way that preserves the basic features of the full sample design. A set of estimation weights (referred to as "replicate weights") will then be constructed for each jackknife replicate. Using the full sample weights and the replicate weights, estimates of any survey statistic can be calculated for the full sample and each of the 50 jackknife replicates. The variability of the replicate estimates is used to obtain a measure of the variance (standard error) of the survey statistic. Previous surveys, using similar sample designs, have yielded relative standard errors (i.e., coefficients of variation) in the range of 2 to 10 percent for most national estimates. Similar results are expected for this survey.

B.3. Methods for Maximizing the Response Rate for the Probe study

As recommended by OMB, an advance letter, with salutation by name, will be mailed to the President of the institution requesting that the President identify an appropriate survey respondent and provide the respondent's contact information (see letter to presidents, Attachment A-1). Based on findings from the feasibility study and Westat's experience in conducting surveys that are nonrecurring or sponsored by agencies that are not well-known by the postsecondary community, it is important that the letter contain an information copy of the questionnaire and an enclosure with information about the Governing Board and its preparedness research program (Attachment A-3). These documents will help the institution to determine whether they would be willing to participate in the survey. The enclosed questionnaire will be clearly stamped as an information copy that should not be completed.

The President's office is not expected to be the location of the appropriate respondent in the vast majority of cases. Early survey development work found and findings from the feasibility study confirmed that the person or persons responsible for evaluating incoming students' need for developmental education can be found in offices such as:

- Offices of academic deans or provosts;
- Academic departments (e.g., math or English departments);
- Offices of institutional research;
- Offices of student assessment;
- Offices of student services
- Offices of student counseling;
- Undergraduate admissions offices;
- Student orientation offices; and
- Corporate offices (primarily in for-profit institutions).

Given this variability, the President's office is a logical starting point for identifying the respondent, since this office will be most familiar with the overall organization of personnel and responsibilities within the institution. This has been confirmed through consultations with NCES, feedback from experts in postsecondary education, and findings from the feasibility study.

Survey packages will be mailed, with salutation by name, to designated survey respondents. The package will contain a cover letter (Attachment A-2) indicating that the designated respondent has been identified as the appropriate survey respondent by the President's office and requesting participation in the survey. The package will also contain an enclosure (Attachment A-3) about the Governing Board and how this survey fits into its overall program of research on the academic preparedness of 12th graders. The cover letter will encourage respondents to complete a web version of the survey, and it will also offer the alternative of completing a traditional paper and pencil questionnaire. The survey is limited to eight questions; these eight questions are covered in four pages of the paper version of the survey.

Telephone follow-up for nonresponse, which will be conducted by Westat staff, will begin about 3 weeks after mail-out for each type of collection. Experienced telephone interviewers will be trained to conduct follow-up and will be monitored by Westat supervisory personnel during all interviewing hours. Collection procedures will follow standard methods developed by Westat for the feasibility study and previous surveys on postsecondary institutions.

B.4. Tests of Procedures and Methods

A pretest of the survey was conducted with seven respondents to determine what problems respondents might have in providing the requested information and to make appropriate changes to the questionnaire. Responses and comments on the questionnaire were collected by fax and telephone during the pretest. The results were used to make revisions to the survey items and instructions, and summarized as part of the documentation for the study.

A feasibility study was conducted with a sample of 120 postsecondary institutions in fall of 2010. This feasibility study was originally designated a “pilot study”; however the term “pilot study” was a misnomer. The feasibility study was aimed at answering a number of specific questions, including:

- Is the use of separate forms for 2-year and 4-year institutions supported?
- Should any of the test lists be revised?
- Are the items being correctly understood?
- Is the President’s office the appropriate place to identify the survey respondent?

As a result of this feasibility study, the use of separate forms for 2-year and 4-year institutions was supported; the test lists have been revised; minor adjustments have been made to ensure that the survey items are correctly understood; and the President’s office was confirmed as the appropriate place to identify the survey respondent. In addition, data checks have been added to the web-version of the survey to serve as prompts to address missing responses, out-of-range responses, and other potential data issues.

A study comprised of cognitive interviews and a usability test will be conducted with nine institutions to assess changes made to the questionnaire as a result of the feasibility study and to assess the data checks developed for the web-based version of the survey. The study will include a range of institution types, including public and private, two-year and four-year institutions.

Cognitive and usability aspects of completing the survey will be examined concurrently as participants proceed through the questionnaire. The cognitive aspect will focus on participants’ interpretation of instructions and definitions, as well as wording of key survey questions, particularly those that were modified after the feasibility study. The usability aspect will focus on participants’ interaction with the web survey, paying particular attention to the efficacy of the web data checks added after the feasibility study. The results will be used to make any final modifications to the survey and the data checks. We will inform OMB of any changes to the surveys that might result from this study.

B.5. Reviewing Statisticians for the Feasibility study and Probe study

Westat’s statisticians, Adam Chu (telephone: 301-251-4326) and Greg Norman were consulted about the statistical aspects of the design.