

### Survey Methodology

During the spring of 2007, the Statistical Research Center (SRC) of the American Institute of Physics (AIP) contracted to conduct the Humanities Indicators Survey. The project was directed by Alice Noble and John Hammer of the American Academy of Arts and Sciences. The survey steering committee included representatives from the American Historical Association, the Modern Language Association, and the American Political Science Association. Other societies participating in the survey included the American Council of Learned Societies, the American Academy of Religion, the Linguistic Society of American, the College Art Association, and the History of Science Society.

*Scope of Work:* The project involved conducting a survey of a sample of departments and programs in seven humanities disciplines (listed below). All activities related to the survey were conducted in consultation with the project directors and the steering committee. SRC agreed to develop a questionnaire in both hardcopy and on-line versions, host the on-line questionnaire on SRC servers, select sample departments and programs for each field, contact department and program heads to encourage them to complete the survey, follow-up with non-respondents with the goal of a 60% response rate, collect data, enter the data into an electronic database, and prepare the data for analysis. The SRC also agreed to provide the project team with a code book that described the content of the electronic database, as well as this report which details how the project was conducted, the problems encountered, and suggestions for changes in how future surveys might be conducted.

*Disciplines:* As a pilot project, The Humanities Departmental Survey limited the participating disciplines to the following seven: history, religion, college art, linguistics, English, foreign languages, and history of science. It was assumed that the contact information maintained by the learned societies that collaborated on this project reflected with reasonable accuracy the population of all departments and programs in their respective disciplines.

History of science was included in the study as a test of whether program heads could provide the detailed information asked for in the questionnaire. Only 20 such programs were included in the sample. History of science is comprised largely of degree-granting programs rather than stand-alone departments. There was concern about whether respondents could provide accurate data on those faculty and students involved in the history of science program as separate from all faculty and students in the department housing the HSS program.

*Criteria for Inclusion:* Several criteria were used to determine whether specific departments and programs qualified for the study. First, departments or programs had to award degrees in one or more of the seven target fields, and the departments or programs had to be in a four-year institution in the United States. The sample was selected so that it would accurately represent degree-granting departments and programs in three types of four-year institutions: research intensive, comprehensive, and primarily undergraduate.

The taxonomy for several disciplines was broad, and field boundaries had to be clearly defined. The project directors and steering committee decided that the survey should focus on scholarly fields. The Indicators Survey intentionally excluded variations of the target fields that were

classified as applied. For example, the American Academy of Religion database included departments that award degrees to people preparing for ministry, and these were excluded from the survey.

### **Population Information**

Each society that was collaborating on this project maintained a membership database that effectively addressed the information needs of the society. The survey methodology relied heavily on these databases to provide a reasonably accurate and complete picture of the population of all degree-granting departments and programs in the target fields. While these society databases were a good starting point, four issues quickly emerged: was the information current; was the information accurate; was the information complete; and were all degree-granting departments and programs represented? Cleaning up the databases was more time consuming for some disciplines than for others. However, this step is absolutely necessary. If a sample is drawn from the wrong population, then it is impossible to generalize the results to the correct population.

*It is also important to note that problems in the databases could have been identified at three points. The first is before the sample was drawn, and details of this process are discussed in this section. The second is during the process of data collection, and that process is discussed in a subsequent session. The third is during data analysis.*

**CAA:** The College Art Association's database included information on which areas of specialization were available in each of the departments. These fine fields included studio art, architecture, film, fashion, and graphic design. However, the project steering committee had previously decided to include only art history departments in the survey. One complication was that CAA's database did not include undergraduate programs. In order to have access to the full population of departments, SRC staff members supplemented the college art database with information from the Integrated Postsecondary Education Data System (IPEDS), which is an information resource developed and maintained by the US Department of Education. IPEDS was also used to verify whether the CAA database covered all departments that awarded graduate degrees in art history. In addition, Internet searches were needed to identify current department chairs and formal department names for those art programs (both undergraduate and graduate) that were selected into the sample, but were missing from the CAA database.

**AHA:** The history database included about 1,000 institutions with history degree-granting departments and the AHA maintained current contact information for nearly 800 of them. AHA staff members conducted Internet searches for the names and contact information of current department heads for the remaining 200 institutions.

**MLA:** The Modern Language Association's list of about 3000 departments included English departments, foreign language departments, and departments that offer courses in both English and foreign languages. The latter are referred to as "combined" departments throughout this report. Although the combined departments offered courses in English and in foreign language, many of them do not offer degrees in both. Fifteen of these were removed before the sample was drawn. MLA also included 90 linguistics departments that were already part of the LSA

database and therefore were removed. At schools that had two departments that offered an English degree, the traditional “English” department was retained and the secondary department removed. For example, 53 stand-alone comparative literature and four English as a second language departments were removed. At schools with two departments offering the same foreign language, efforts were made to select the main department where the language was taught.

**LSA:** According to the information the SRC received from the LSA, about 40% of their 176 programs had the word “linguistics” in the department name. About one-fourth were linguistics programs that reside in English or literature departments. The rest were programs that reside in many different departments, including various foreign languages, cognitive science, and anthropology.

**AAR:** The religion database had 1,818 departments and schools that offer degrees according to the AAR. In order to remove programs whose purpose was the training of ministers, priests, and rabbis, SRC staff members wrote software to remove nearly 750 institutions and departments with words like ministry, mission, music, seminary, and theology in their names. SRC staff members removed 157 departments that were outside the US. The IPEDS database did not contain 14 AAR programs, which indicates that they did not actually award bachelor’s degrees. These were removed. Staff members also deleted more than 40 schools that were classified as granting an associate’s degree, as well as nearly 60 departments that had duplicate records in the database. Most of the latter were departments that had both an undergraduate and graduate listing. After these deletions, the database still had institutions with several departments. In these cases, we kept only the department called “religion”. The tally of qualified departments from which we drew the sample was about 600.

**HSS:** SRC staff members were concerned about the History of Science list on two levels. First was the small number of programs in the HSS database, because this limits analytical opportunities. Second, virtually all of the contacts were heads of programs rather than departments of the history of science. SRC staff members were uncertain if the program contacts would be able to provide data on the program only. SRC staff members recommended that the study include larger history of science programs and, using data from NSF and Department of Education, staff members identified the 20 institutions that awarded at least 4 graduate degrees in the history of science *and technology* over the last 5 academic years combined. History of science cannot be disaggregated from the history of technology in the federal databases. We consulted with the History of Science Society’s director, who agreed that history of technology programs could be included as well.

### **Sample Selection**

In order to achieve the analytical goals of the study and provide sufficient data for reliable results, the SRC considered several different sampling strategies. Discussions with the steering committee made it clear that one essential criterion was that the analyses highlight the differences across the Carnegie classification of the institutions. The data the SRC received from the societies that were collaborating on this study showed that the number of degree-granting programs and departments varied dramatically across disciplines. SRC staff members concluded that a strategy that would provide reliable data for each of the target disciplines would be to sample departments and programs within each field separately and that these discipline-specific

## The 2007-08 Humanities Departmental Survey

samples would be stratified by three levels of Carnegie classification: research intensive, comprehensive, and primarily undergraduate.

SRC staff members took these dynamics into account and assumed a response rate of 60% in calculating the minimum sample size that would provide reliable statistics. We recommended that the sample be comprised of 80 departments or programs within each field from each of three general Carnegie classifications. When the population of departments in a discipline at any one of the three types of universities was close to 80 (e.g. art history at bachelors-granting colleges), then we selected all departments in that category.

Table A1 shows the number of departments and programs that were estimated to be in the population for each field (other than history of science) at each of the three types of institutions. The bulk of this information came from the databases provided by the disciplinary societies. This information was augmented by information from IPEDS and Internet searches in the following cases: undergraduate programs in art history, the approximately 500 history departments with which AHA did not correspond regularly, and the history of science sample. Table A1 also shows the number of departments and programs that were drawn into the sample from each category of field and institution. The number of programs that awarded degrees in linguistics was small enough that all of them were included in the study. Programs in the history of science are not included in Table A1. These were selected separately and represented the 20 largest degree-granting graduate programs based on data collected by the National Science Foundation.

### **Questionnaire Development**

In consultation with the survey steering committee, the SRC revised the questionnaire during the summer of 2007.

The questionnaire was pretested on-line during late October and early November 2007. After we drew the sample, we selected departments that were not included in the sample for the pretest. The pretest was conducted one discipline at a time so that it would be easier to spot wording issues that might be field specific. The dates for these pretests were:

- 31 history department chairs were asked to participate on Oct. 15;
- 30 religion department chairs were asked to complete the on-line questionnaire on Oct. 26;
- 15 art history department chairs were asked to complete the questionnaire on Nov. 1;
- 30 MLA English department chairs were asked to participate on Nov 2;
- 6 MLA foreign language chairs were pretested on Nov. 8.
- An additional 29 history department chairs were asked to participate in the final pretest after the on-line questionnaire had been modified to address wording problems that emerged from earlier pretests.

The 2007-08 Humanities Departmental Survey

**Table A1: Original Population Estimates and Sample Sizes**

Discipline	Population N	Sample N
<b>History</b>		
Bachelors	285	80
Masters	413	80
Doctoral	237	80
Subtotal	935	240
<b>Religion</b>		
Bachelors	284	80
Masters	196	80
Doctoral	123	80
Subtotal	603	240
<b>MLA English</b>		
Bachelors	414	80
Masters	470	80
Doctoral	235	80
Subtotal	1,119	240
<b>MLA Foreign Language</b>		
Bachelors	470	80
Masters	392	80
Doctoral	577	80
Subtotal	1,439	240
<b>MLA Combined</b>		
Bachelors	63	36
Masters	94	36
Doctoral	6	6
Subtotal	163	78
<b>Art History</b>		
Bachelors	101	101
Masters	90	90
Doctoral	148	80
Subtotal	339	271
Linguistics	176	176

No department chairs or programs heads from HSS, MLA combined, or linguistics were in the pretests since these groups were so small that we needed all of them for the full study. During the pretests, a number of problems in the wording were identified.

After the pretests, we developed the paper version of the questionnaire. The on-line version was somewhat different from the paper version, because the on-line version allowed respondents to automatically skip blocks of questions that were not applicable.

### **How the Survey Was Conducted**

The survey sample was comprised of a mix of departments, programs within larger departments and programs that resided in the space between departments. It was essential that each potential respondent be given clear directions identifying the target discipline about which he or she was being asked to report. In other words, program heads needed to know that we were asking them about their program and not about the entire department. Clearly, this issue affected some disciplines like history of science and linguistics far more than other disciplines.

Target disciplines were identified for the respondents in two ways. First, cover letters, whether paper or electronic, were addressed to specific individuals, and each letter identified the specific target discipline about which that individual was being asked to report.

Second, the target field was identified on the questionnaire. For the paper version, a label was printed and affixed to the lower right-hand corner of the front of the questionnaire. Printed on the label was the ID number that corresponded to the department that was receiving it, as well as the phrase "Please respond about <name of discipline inserted here>". This procedure required considerable coordination during the preparation of the mailings, as the outside envelope, the cover letter, and the questionnaire with label affixed needed to be matched for each of the 1,505 programs in the sample.

The electronic equivalent of these procedures was used for the on-line version of the survey. Each e-mail was addressed to a specific person and identified the target field about which the individual was being asked to respond. In addition, the e-mail contained a personalized link to the on-line survey. Once the respondent clicked the link, our software inserted the name of the discipline in appropriate questions throughout the on-line questionnaire. The name of the discipline appeared on the introductory page of the on-line survey in the phrase "Please respond about <name of target discipline>". The software also inserted the field name in many questions, such as "How many of the FULL-TIME TENURED faculty members employed in <name of target discipline> were MEN?"

Although the data collection phase of the study was comprised of both paper and on-line questionnaires, the paper version was viewed as the primary mode. One week prior to the first mailing of the paper questionnaire, department chairs and program contacts were sent a pre-notification letter explaining who was conducting the study and why it was important. This letter was sent by the SRC on the appropriate society's letterhead and was signed by the society's executive director. The pre-notice included a five-page document prepared by the steering

## The 2007-08 Humanities Departmental Survey

committee explaining the study in more detail. About one week after the first mailing of the paper questionnaire, a postcard reminder was mailed to contacts in all of the departments and programs in the sample.

Non-respondents were sent reminders encouraging them to respond. All contacts, except for the final contact, were sent by the SRC but appeared to come from the executive director of the appropriate society. In all, there were a pre-notice, a postcard, three mailings of the paper questionnaire, and three e-mail reminders. The final contact, which was the third mailing of the paper questionnaire, was sent by certified mail in order to emphasize its importance to non-respondents. This mailing was sent by the SRC, but was signed by the Academy and by the American Council of Learned Societies (ACLS). This switch was made in the hopes that departments that had not responded to requests from the societies' executive directors would respond to a request from the Academy and ACLS. The final contact had very limited success, but since these were respondents who had ignored the first six requests, the low success rate cannot be attributed to the change in the signatories. See Appendix C for the schedule of mail and e-mail contacts for each target discipline.

Undeliverable mail was returned to the SRC, and we also received e-mail correspondence from individuals who indicated that they were no longer the department chairs. Internet searches were conducted to identify 71 new department chairs. As Appendix C indicates, the new department chairs were contacted on a separate schedule beginning in February 2008.

Some respondents contacted us to indicate that their department or program does not offer a degree in the target discipline. These programs were removed. In January, we looked through the paper returns for the linguistics programs and became concerned that some of the programs did not offer degrees in linguistics. We checked all of the linguistics non-respondents at that time and removed 36 of the original 176 linguistics programs. Across all fields, 88 departments and programs were deleted from the study during data collection. The sample size after removals is indicated in the second column of Table A2.

SRC staff members concluded that the 88 removals were representative of a larger number of departments and programs that were incorrectly included in the original population estimates for each target discipline. Thus, we did not replace the departments that were removed. The sample for history had the fewest deletions, in part, because 60% of the sample came from a well-maintained society database, and 40% came from searches of IPEDS and the Internet by SRC staff members prior to the first mailing.

During the data collection phase, 48 department chairs and program heads (about 3% of the sample) declined to participate and asked to be removed from the study (Table 2). Overall, 66% of the departments and programs responded to the questionnaire. History departments had the highest level of cooperation, with a 73% response rate. The combined English and foreign language departments had the lowest response rate (60%), but the combined departments were a very small group.

**Table A2: Original and Adjusted Sample Sizes with Field-Level Response Rates**

Discipline	Sample N	Sample N After deletions	Responses N Paper	Responses N On-line	Resp Rate	Refusals N
<b>History</b>						
Bachelors	80					
Masters	80	238	103	69	72%	7
Doctoral	80					
<b>Religion</b>						
Bachelors	80					
Masters	80	215	90	47	64%	10
Doctoral	80					
<b>MLA English</b>						
Bachelors	80					
Masters	80	236	81	67	63%	9
Doctoral	80					
<b>MLA Foreign Language</b>						
Bachelors	80	232	111	52	70%	2
Masters	80					
Doctoral						
<b>MLA Combined</b>						
Bachelors	36					
Masters	36	75	25	21	61%	7
Doctoral	6					
<b>Art History</b>						
Bachelors	101					
Masters	90	262	128	48	67%	9
Doctoral	80					
Linguistics	176	140	53	36	64%	4
History of Science	20	19	4	9	68%	0
Totals	1505	1417	595	349	67%	48

**Calculation of Population Estimates**

Responses from the sampled departments are used to calculate the population estimates. The data are weighted by field and by Carnegie classification. When estimating the total number of “variable of interest,” the following formula is used:

## The 2007-08 Humanities Departmental Survey

$$\hat{x}_{fq} = \sum_{c=1}^3 w_{fc} * rr_{fqc} * x_{fqc}$$

where  $\hat{x}_{fq}$  denotes the total population estimate for a particular field ( $f$ ) based on responses to a particular question ( $q$ ),  $w_{fc}$  denotes the weight associated with that field for that Carnegie classification ( $c$ ),  $rr_{fqc}$  denotes the adjustment to the weight for item non-respondents for that particular question, field, and Carnegie classification, and  $x_{fqc}$  denotes the sum of the sample data for that field for that question for that Carnegie classification. The individual weights ( $w_{fc}$ ) are determined by comparing the total number of departments in the population to the total number of departments in the sample, and the item non-response adjustment ( $rr_{fqc}$ ) is determined by comparing the total number of responses to a particular question from respondents in a particular field and Carnegie classification to the total number of respondents in that field and Carnegie classification overall.

When the data are presented as proportions, the totals were calculated as described above first, and those totals were used to calculate the proportions.

### **Problems with Specific Questions**

A preliminary examination of the data indicated that most of the questions appear to work and that respondents seemed to be providing the data the questionnaire asked for. There were a few significant exceptions, all of which were on the third page of the paper questionnaire.

*Interdisciplinary Programs:* We strongly suspect that most of the data from this section will not be usable. One of the questions asked respondents to list the names of the interdisciplinary programs with which their department was involved. Some department heads responded about specific courses even though the question asked about programs. However, the disciplinary societies may be able to glean some useful information from some of the answers written in response to this question.

Respondents also had trouble with the question reading “How many students participated in these [interdisciplinary] programs during the 2006-07 academic year?” The definition of students was not clear (students who are majors in the target discipline? general students?) and the definition of participated was not clear (taking classes? majoring?).

*General education courses:* This is a term that appeared to have different meanings to respondents in different universities. Many respondents apparently thought that they had no general education courses and left this set of questions blank. Many respondents could not distinguish between general education and lower division courses. These individuals provided identical data for the two sets of questions. There were so many obviously incorrect answers to the set of questions about general education courses that we recommend that the data be ignored.

*Lower division courses:* This term appears to have created some of the same confusion as general education courses. We believe, however, that the bigger problem with the set of questions about lower division courses may have resulted from a sentence that the steering

## The 2007-08 Humanities Departmental Survey

committee added to the question about lower division courses. This sentence defined lower division courses as courses that were taken by undergraduates who intended to major in the target discipline. Many respondents wrote that they could not distinguish the intended majors from the total enrollment in lower division courses. Different respondents dealt with these problems in different ways, but there were so many obviously incorrect answers to the set of questions about lower division courses that we recommend that these data be ignored.

*Number of sections:* The layout of the questions about the number of sections appears to have contributed to some confusion as well. These questions asked about the number of sections taught by faculty in lower division, general education, upper division, and graduate levels. However, a few respondents did not read the question or, at least, did not read it carefully. For example, some respondents appeared to tell us how many faculty members were teaching sections instead of how many sections were offered.

*Number of faculty:* Several program heads did not complete the questions that asked about the number of faculty by tenure status and gender. Fortunately, some respondents took the time to write comments about this set of questions in the open-ended question at the end of the survey form. Many of these respondents indicated that their programs existed in the space between formal departments and that these programs had no dedicated faculty. Whether the heads of degree-granting programs can provide counts of their faculty is an especially difficult problem for interdisciplinary programs. This problem should be addressed in future surveys of programs.

### **Conclusions**

This survey asked chairs of a sample of departments in seven fields information about their faculty, students, courses, hiring, turnover, and policies. The fields included were history, religion, art history, English, foreign languages, linguistics, and history of science. The survey was administered both on the web and in paper from November 2007 to May 2008. The overall response rate was 66%.

This survey relied on participating societies to provide complete lists of degree-granting departments and programs, as well as current and accurate contact information. However, societies keep lists of departments for their own purposes. For example, many departments and programs appeared on societies' lists but were in fact ineligible for the study because they did not offer a degree or were only a part of a larger, degree-granting department. The lists contained duplicates which are easily removed, but the inclusion of many programs that do not offer degrees is especially problematic. Also problematic is the omission of programs that do offer degrees. In order to generalize to the larger population, the lists must be cleaned before the sample can even be drawn. In addition, the societies' lists contained outdated contact information, a problem that will exist with any list and is time-consuming to correct, but does not have serious implications for sample selection.

This survey was designed as a template and served that purpose well. Although the questionnaire was pretested, problems with several questions were discovered during and after data collection. Future rounds of this survey will provide opportunities to refine the questions.

**Recommendations for changes in how future surveys might be conducted**

This survey used a single instrument for departments, for programs within departments, and for programs that draw faculty members from several departments. Several program heads noted their difficulty with answering the questions about the number of faculty, but it is unknown how many answered these questions in a way that does not give the answers desired by the committee. It is possible that in the future, separate questionnaires will be needed for departments and programs. This will increase the complexity of the administration of the survey and its expense.

The process of adding new programs should be undertaken slowly and with caution. Until a process for collecting accurate data from programs is developed, there is always a risk of over-counting students and faculty members. For example, faculty members teaching in a program may be counted twice—once by their home department and again by the program. It is also not clear what the concept of a lower division course means to a specialized program such as history of science. Will these lower division courses simply be duplicates of the courses offered by the history department?

This survey relied on communications that appeared to come from the executive directors of the participating societies, except the final paper mailing which was signed by the Academy and the ACLS. Letters coming from executive directors appeared to be effective, but increase the complexity and expense of survey administration. Future rounds of this survey should attempt to determine whether letters coming from the Academy and ACLS would be just as effective. The SRC continues to recommend that future rounds of this survey continue to utilize paper questionnaires, which are preferred by many respondents when questions ask for departmental records.

It is our understanding that the number of disciplines and interdisciplinary programs will increase in future rounds of the survey. SRC staff members recommend that the same 1417 departments and programs be surveyed again in the next round. This would simplify the process of data collection so that efforts can be expended on identifying the boundaries of the new fields.



### **Questions That Did Not Work**

As with any initial attempt, the first Humanities Departmental Survey offers many opportunities to learn lessons. This section addresses the lessons to be learned from questions that did not work on the survey instrument.

#### **Interdisciplinary Programs**

The series of questions about interdisciplinary programs proved challenging. Perhaps some of the respondents are housed in interdisciplinary programs, so, in the next round, one of the answer choices could reflect this reality. In response to the question about the names of interdisciplinary programs, some participants appeared to provide a list of classes rather than the name of a program. Finally, it is possible that respondents are not responsible for tracking enrollments in all of the interdisciplinary programs in which they participate. So, it is not clear that this data is useful.

#### **Undergraduate General Education Courses**

##### **Undergraduate Lower Division Courses**

The questionnaire included a set of questions about general education courses and a set of questions about lower division courses. These questions asked about the total number of students enrolled during the fall of 2007, the total number of sections in these course, and who taught those section, i.e. full-time tenured faculty, full-time non-tenured faculty, part-time faculty, and graduate students.

These questions were intended to provide fundamentally important indicators of the vitality of the department or program. However, these two sets of questions created so much confusion among respondents that it is our opinion that the data are neither reliable nor accurate.

Some of the confusion centered on the distinction between general education courses and lower division courses. An example of this problem is that 137 of the 944 respondents wrote the same exact answers to both sets of questions. In addition, based on comments written by a number of respondents, some people were confused by the definition of lower division courses which stated that these courses were primarily intended for students “who might major in the field.” Examples of these problems are reflected in the 117 respondents left the lower division course information blank, despite the fact that virtually all of them reported that they offered an undergraduate major. The level of confusion is further indicated by the 111 respondents who simply left both sets of questions blank.

In short, fewer than half of respondents answered both sets of questions providing different answers to the two sets.

Of the respondents who did answer these two sets of questions, many provided inconsistent answers to the number of sections taught. Unfortunately, these problems appeared in many ways and in the responses to both the general education courses and the lower division courses.

One problem was that many respondents appeared to be giving us data on the number of individual “discussion” sections and similar classes taught by graduate students as part of a

larger course. When we added the number of sections reportedly taught by graduate students to the number taught by faculty, the sum far exceeded the answer to the “total number of sections” on their questionnaires.

On several dozen questionnaires, we had the opposite problem. When we added together the number of sections taught by tenured faculty, non-tenured faculty and graduate students, we had a sum that was far smaller than the answer these respondents provided to the “total number of sections taught” in the same course.

Another problem was that some respondents appeared to be giving us data on the number of students in the sections rather than data on the number of sections taught by graduate students and faculty. In such cases the sum totals for these answers were close to the answer to “how many students were enrolled in” the course and far exceeded the total number of sections in the courses.

Finally, we could not deduce what questions several dozen respondents thought that they were answering. When we added their answers to the number of sections taught by tenured faculty, non-tenured faculty, and graduate students, we had a sum that was more than 100 greater than their answer to the “total number of sections,” but several thousand smaller than the total number of students enrolled in the course.

There is also potential confusion between “students” and “enrolled.” It is possible that some smaller departments counted individuals and larger departments counted enrollments. A smaller department might know that it has six majors with each one enrolling in three different courses. So, is the total number of students 3, or is the total enrollment 18? This potential for confusion runs throughout all of the questions about classes and independent study.

### **Graduate Level Independent Study**

For some disciplines the sum of students taught by the three types of faculty members was very different (by as much as 45%) from the number reported in “Total number of students enrolled in graduate level independent study.” It could be that some respondents differentiated between “enrollments” and “students.” If a single individual is doing a graduate level independent study with several different faculty members, it is not clear how that should be reported. Is the goal here to determine how many different “projects” are underway (with one project potentially including multiple faculty members) or how many different “registrations” exist within a department during a given semester.

### **Union Representation and Benefits**

These questions apparently caused both survey fatigue and respondent confusion. For these questions, we first examined responses from departments at the same school. We looked at answers to questions 38 – 42 for 525 respondents from the same school. There are 150 pairs with two departments from the same parent institution (300 respondents), 45 triads with three departments from the same parent institution (135 respondents), 12 instances with four departments from the same parent institution (48 respondents), 6 instances with five departments from the same parent institution (30 respondents), and 3 instances with six departments from the same parent institution (12 respondents).

## The 2007-08 Humanities Departmental Survey

For these five questions, 15% of the responses were missing, and about 40% of the non-missing responses were inconsistent within the same school. In There are ten answers required of respondents for these five questions (two answers for each question, one for full-time faculty and one for part-time faculty). In only three of the ten cases were more than half of the responses non-missing and consistent. The pattern of inconsistencies was not consistent. For question 38, the responses were more consistent within full-time faculty; for questions 39, 40 and 42, the responses were more consistent within part-time faculty; and, for question 41, the consistency (or lack thereof) was very similar for both full- and part-time faculty.

Both the high level of omission (15% of the responses were blank) and the high proportion of inconsistent responses (40% of the non-missing answers are inconsistent within the same institution) suggest that departmental personnel are not necessarily the best respondents for these questions. Furthermore, asking these questions multiple times within the same institution requires duplication of effort on the part of the respondent.

We recommend asking these questions once for each institution and asking the Human Resources office to respond, not the departments. It is quite possible that the person completing the form at the departmental level is not well-versed in benefits for faculty members, particularly when looking at fine distinctions between full-time and part-time faculty.

The definition of part-time faculty is not clear. It is possible that some part-time faculty are benefit-eligible, while others are not.

The answer choices should be mutually exclusive. The choices used for many of the benefits questions on the 2007 survey were:

- Yes, fully paid by institution
- Yes, partially paid by institution
- Yes, paid by the individual
- Not available
- Don't know

We suggest the following choices instead:

- Yes, fully paid by institution
- Yes, partially paid by institution
- Yes, fully paid by individual
- Not eligible for plan
- Not available
- Don't know

This revised list clears up any confusion about which response is best when the costs are shared between the employee and the employer. In addition, it provides for a distinction between a particular benefit not being available at an institution at all versus the particular class of employee not qualifying for the benefit; in the former list it is not clear which is meant by a response of "not available."

## The 2007-08 Humanities Departmental Survey

Finally, if the question regarding criteria used to determine eligibility for part-time faculty and instructional staff were asked before the questions about benefits, it might be clearer to the respondent what is meant by “part-time faculty.”

### Sampling Error and Confidence Intervals

Sampling error is the error caused by using data from a sample (rather than data from a census of the whole population). All of the population estimates in this report are subject to sample-to-sample variation; that is, responses from a different sample of survey participants could (and likely would) result in slightly different estimates. A large enough random sample results in sampling errors that are relatively minor. We can estimate the magnitude of the sampling errors using standard statistical formulas.

For data that are presented as proportions, the estimated size of the sampling error varies with the magnitude of the particular proportion in question and the size of the sample (or sub-sample) in consideration. It is given by:

$$s = \sqrt{\frac{\hat{p} * (1 - \hat{p})}{n}}$$

where  $\hat{p}$  is the estimated proportion and  $n$  is the number of observations in the sample. So, for example, we estimate that 77% of the departments in Art History participate in interdisciplinary programs (Table 13). There were 168 respondents from Art History departments who answered the question we used to make this estimate. So, the sampling error for the estimated proportion of Art History departments participating in interdisciplinary programs is:

$$s = \sqrt{\frac{0.77 * (1 - 0.77)}{168}} = 0.0325$$

A confidence interval provides an interval estimate of a population parameter. The width of the confidence interval indicates the reliability of the estimate. A confidence interval is given by:

$$\hat{\theta} \pm Z * s$$

where  $\hat{\theta}$  is the point estimate (such as the numbers given in tables throughout this report),  $Z$  is the confidence coefficient, and  $s$  is the sampling error. The confidence coefficient for a 95% confidence interval is 1.96. So, the 95% confidence interval for the proportion of Art History departments that participate in interdisciplinary programs is:

$$0.77 \pm 1.96 * 0.0325 = 0.77 \pm 0.06 = \{0.71, 0.83\}$$

The level of confidence is associated with the estimation process itself and is unrelated to the point estimate.

The 2007-08 Humanities Departmental Survey

For data that are presented as population totals derived from sample data, the size of the sampling error varies with the weighting factor<sup>18</sup>. It is calculated as:

$$s = \sqrt{w^2 * \hat{\sigma}^2}$$

where  $\hat{\sigma}$  is the sample standard deviation and  $w$  is the weighting factor.

Below we provide the estimated standard error and 95% confidence intervals for the total faculty data (Table SE1) and for the total number of bachelors degrees completed (Table SE2). In most cases, the result for a 95% confidence interval ( $1.96 * s$ ) is about 2 to 4% of the estimate. The variations in the relative size of the sampling error are as expected based on differences in sample sizes and standard deviations for each sample.

**Table SE1: Estimated Sampling Error and 95% Confidence Intervals for Total Number of Faculty Members by Carnegie Classification**

Discipline	Carnegie Classification	Estimated Total Faculty	Estimated Sampling Error	Lower Bound, 95% CI	Upper Bound, 95% CI	Approximate Range
Art History	Bacc. Coll.	490	5.0	480.1	499.9	±2.0%
	Masters U.	560	6.8	546.7	573.3	±2.4%
	Doc. U.	1,750	14.7	1,721.1	1,778.9	±1.7%
English	Bacc. Coll.	4,910	78.7	4,755.7	5,064.3	±3.1%
	Masters U.	11,970	183.4	11,610.5	12,329.5	±3.0%
	Doc. U.	13,800	144.2	13,517.4	14,082.6	±2.0%
Foreign Languages	Bacc. Coll.	3,530	46.9	3,438.1	3,621.9	±2.6%
	Masters U.	5,750	80.4	5,592.5	5,907.5	±2.7%
	Doc. U.	14,040	199.1	13,649.7	14,430.3	±2.8%
History	Bacc. Coll.	2,540	23.3	2,494.4	2,585.6	±1.8%
	Masters U.	6,180	86.6	6,010.2	6,349.8	±2.7%
	Doc U.	6,640	61.6	6,519.3	6,760.7	±1.8%
Linguistics		1,630	14.1	1,602.4	1,657.6	±1.7%
MLA Combined	Bacc. Coll.	1,100	24.2	1,052.6	1,147.4	±4.3%
	Grad. U.	2,270	55.5	2,161.2	2,378.8	±4.8%
Religion	Bacc. Coll.	1,820	27.3	1,766.5	1,873.5	±2.9%
	Masters U.	1,710	27.2	1,656.7	1,763.3	±3.1%
	Doc. U.	1,480	20.7	1,439.3	1,520.7	±2.7%

<sup>18</sup> Since the weighting factor varies inversely with the sample size, the sampling error decreases as the sample size increases.

**Table SE2: Estimated Sampling Error and 95% Confidence Intervals for Total Number of Bachelors Degrees Awarded by Discipline**

<b>Discipline</b>	<b>Estimated Number of Bachelors Degrees Awarded</b>	<b>Estimated Sampling Error</b>	<b>Lower Bound, 95% CI</b>	<b>Upper Bound, 95% CI</b>	<b>Approximate Range</b>
<b>Art History</b>	5,400	32.2	5,336.9	5,463.1	±1.2%
<b>English</b>	54,690	624.5	53,466.0	55,914.0	±2.2%
<b>Foreign Languages</b>	28,710	186.7	28,344.1	29,075.9	±1.3%
<b>History</b>	38,700	279.1	38,153.0	39,247.0	±1.4%
<b>Linguistics</b>	1,720	26.85	1,667.4	1,772.6	±3.1%
<b>MLA Combined</b>	2,980	63.5	2,855.5	3,104.5	±4.2%
<b>Religion</b>	5,160	43.1	5,075.5	5,244.5	±1.6%

The estimated standard errors are smaller, in relative terms, than those for the faculty data presented in Table SE1. This is because the data in Table SE2 are aggregated into broader categories than those in Table SE1.

### **Other Sources of Error**

Many sources of error arise in survey research. These include:

- Non-response bias which results when there are systematic differences in response characteristics between those who responded to the survey and those who did not respond; all of the estimates are based upon respondents who may not necessarily be entirely representative of the population
- Errors arising from poorly worded questionnaire items or from poorly constructed or unduly complex questions
- Errors in the interpretation of the questions or recall of the responses by respondents
- Errors in data entry and in statistical computation

Every reasonable effort has been made to minimize errors at every stage in this research effort.