

STUDENT EXPERIENCE IN THE RESEARCH UNIVERSITY (SERU) PROJECT

CSHE | Center for Studies in Higher Education

UNIVERSITY OF CALIFORNIA, BERKELEY

<http://cshe.berkeley.edu/>

SERU Project Technical Report¹

Listening to Students' Advice— An Analyst's Story About the Results and Challenges of Using Open-Ended Items on a Large Scale²

DRAFT JAD x.xx.07 November 2007

Steve Chatman

University of California, Berkeley

Copyright 2007 Steve Chatman, all rights reserved.

ABSTRACT

In the spring of 2006, nearly 60,000 undergraduates across the state responded to the University of California Undergraduate Experience Survey. They were asked, "What are the three most important things your campus could realistically do to create a better undergraduate experience for students like you?" About 45,000 gave written advice. Among the questions that will be addressed in this paper are the following. What advice do students offer and does it vary by student characteristics? How can individual researchers process the huge amounts of qualitative data? And, how can we both collapse comments into categories to do research and keep the richness of the original comments?

Author's Note

Inspired by Ira Glass³ and this subject to offer a differently structured paper, this document is organized in four acts. Act 1 is titled, *Thank you Monica Lopez*, and follows the discovery and use of an earlier comprehensive qualitative data analysis of

¹ The SERU Project is a collaborative effort based at the Center for Studies in Higher Education at UC Berkeley and focused on developing new types of data and innovative policy relevant scholarly analyses on the academic and civic experiences of students at major research universities. The primary data collection activity of the SERU Project has been the development and administration of the University of California Undergraduate Experience Survey (UCUES).

² This paper was presented as part of California Association for Institutional Research session, Listening to Students' Advice—Results, Practical Challenges, and Legal Issues When Using Open-Ended Items on Large Scale Surveys with Dr. Samuel Agronow, Associate Director of SAS-Admissions & Outreach, University of California, Office of the President.

³ Ira Glass is a public speaker, host and producer of *This American Life*, a Public Radio International (PRI) program.

undergraduate students' responses to a similar question on an earlier UCUES questionnaire. Act 1 is about Ms. Lopez's development of a coding structure, the author's discovery of that structure, and the resulting exuberance that the current project would be successful that caused the author to submit a CAIR proposal. Act 2 is titled, *All Hope is Lost (UCUES 2006)*, and reflects the author's growing awareness that the project would not and probably could not succeed as planned. In Act 3, *If You Only Have a Hammer, Then Every Problem Looks Like a Nail*, the author reverts to investigatory tools and techniques that have served him well over the years but don't really fit the task at hand. Act 4, *Skinning Cats*, describes discovery of a mixed analytical and reporting strategy that is, if not perfect, at least a useful solution. Findings illustrating application of this method are shared but the paper is first about research methods and strategies and specific findings are secondary.

Act 1: Thank you, Monica Lopez

In the summer of 2005, Monica Lopez, a graduate assistant working with the Office for Student Research, took upon herself a daunting task. She would read and categorize thousands of student comments on a 2003 survey. The 2003 survey was of a large random sample of undergraduate students (n=16,000) across the University. Among many other items, students were asked two relatively simple and perhaps naive questions:⁴ "What could the university have done to make your undergraduate experience better up to this point?"; and, "What would you have done differently, if anything, to make your undergraduate experience better up to this point?" Working in a windowless central room where the unit kept printers, servers, office supplies, and graduate research assistants, Monica began the task of reading, interpreting, sorting, categorizing and reporting on the advice of over 4,500 students.

As is typical of these tasks, the process was not linear but more like that described by the adage "Two steps forward, one step back" and it is the "one step back" that makes qualitative text analysis such a frightening prospect to researchers. During the first few days of sorting, the occasional remark that didn't fit into any existing category or the category that is obviously too inclusive because it is coming to dominate the distribution, did not bother her much even though realization meant that she had to go back and reread dozens of comments to sort them into newly formed categories. Not such a discouraging event on day 3 or even day 5, but when day 20 arrives with a new unwelcome growing recognition, the step back must have been very painful. I like to imagine her peacefully, resolutely moving through a huge stack of paper, even though the material was, of course, electronic.

Monica worked diligently through the summer, finished the task, and returned to her life at summer's end leaving behind a report, *Advice from Students on Improving Their Undergraduate Experience at the University of California*. She reported her coding structure, offered major findings and then, as far as I know, she disappeared never to be heard from again. I feel that I have come to know her, if not well, then at least a little, because I have used her words, her thinking about what students meant, the fruits of her difficult summer. It was Monica who determined that the advice given by these young

⁴ University of California Undergraduate Experiences Survey of 2003 sample frame was a random sample of undergraduates attending the eight campuses with undergraduate programs. The survey is Internet-based and most of the communication is by email.

scholars could be placed into one of about 80 cubbyholes that in turn could be sorted into 21 larger categories, and 3 broad areas.⁵ The broad areas and categories were:

- Academic (Advising, Classes, Faculty, Instruction, Climate, Requirements, Resources, and Teaching Assistants)
- Amount, Administration & Management of Services (Delivery of General Services and Services Other than Academic, Financial Aid, Housing and On-Campus Food Services, Health Services and Mental Health Services, Information Availability, Parking, and University Provided Transportation)
- General Campus (Employment Opportunities, Facilities--Social Centers, Outside Commercial, Amount Construction, Fees and Use of Fees--Required & Cost of Attendance, Housing Off Campus, Campus Climate (Nonacademic), and Transportation--Not University)

Before sharing what Monica found, I have two confessions to make. First, I am ignoring the work that she did to understand comments to a later question about the undergraduate experience. Second, I could not help but change her organization in what I consider to be minor but helpful ways. Presumptuous, I know, but there you have it. There is no way to be completely objective about a subjective task.

Monica offered the following summary statements (p. 6).

Academic

- Academics were of particular concern to most students. 74% of the students who responded to this open-ended question commented on at least one aspect of their academic experience that could have been improved.
- 29% were concerned about the availability, access, and/or size of undergraduate classes.
- At 17%, the second most frequently mentioned items concerned faculty (competency, accessibility, and/or commitment) and academic advising (availability, accessibility, and quality).⁶
- Other academic areas mentioned by 5% or more of students were quality of instruction (9%) overall competency of graduate student instructors (7%).

Nonacademic

- Non-academic aspects of the college experience were important to students but less salient. Only 36% of respondents commented on non-academic aspects of their undergraduate experience that could have been improved. Common areas of concern included financial aid, housing, campus resources, and social climate.
- The most frequently cited non-academic item was campus climate (16%). This category included comments regarding the presence or absence of school spirit, the orientation of the campus (student-oriented vs. impersonal), the level of social integration, and racial and ethnic diversity.

⁵ The coding structure is shown in Appendix 1.

⁶

Among trends noted by Ms. Lopez were that concern about advising and campus climate increased with tenure at the university. The most dramatic change with class level was mention of campus climate issues--from 35% of first-year native students' comments to about 50% of seniors.

I was very optimistic at this stage. After all, I had in hand a very reasonable looking coding structure, and I firmly believed there must be a way to automate the coding process. Riding this wave of good spirit I submitted a CAIR proposal.

Act 2: All Hope is Lost (UCUES 2006)

Rewind from today to December of 2005 for a meeting in the library of the Center for Studies in Higher Education where institutional researchers representing the campuses of the system are gathered to craft the next UCUES questionnaire. When the topic of open-ended items arises, the group soon agrees upon a similarly simple and perhaps naive question, "What are the three most important things your campus could realistically do to create a better undergraduate experience for students like you?" It is hard to imagine what we were thinking. It was a moment of madness, hubris and hysteria. Because not only were we increasing the task that took a graduate student most of a summer to accomplish by a factor of three, we fully intended to ask the question of a group ten times as large. That is three times 10 times two months work, or 60 months. Sixty months is five years, so with a two-year survey cycle, analysis would be an impossible task for a full-time analyst and one that would appear increasingly hopeless over time.

What was this group of experienced institutional researchers thinking? It is hard to know precisely, but I think that we can identify at least three lines of reasoning. First, they thought that they had to ask for three comments because if they asked for only one everyone would write about parking, financial aid, or dorm food. By asking for three things, they hoped to get past the more mundane and gather pearls of wisdom. Second, they believed in the miracle of technology. They believed that there was a silver bullet, killer app, that would do most of the work. Oh sure, there would be some work, but in for a penny, in for a pound. It wouldn't be that much worse to perform the task for nine campuses then it would be for one. And that brings us to probably the most important reason of all. Surely someone else would do the work. For eleven of the twelve people around the table that was true, and for the twelfth, me, it was almost true. I would acquire the silver bullet, killer app, and I would find a part-time worker to do the heavy lifting. The questionnaire was finished, about 150,000 students were invited to participate, and about 38% did. Nearly 45,000 students offered written advice.

Now come forward in time to the summer of 2007. It's almost two years after the library meeting and the tens of thousands of students' observational nuggets continue to gather dust even as plans were being made to collect another mountain of data. Having made no serious assault on the 2006 mountain and having not read the comments raises some interesting questions about confidentiality, obligation, and liability. But for now, let's just say that guilt and fear were sufficient motivators to try and do something useful with the data in hand. Guilt that students had offered carefully considered advice that no one, not even a powerless student worker, had read. Fear that another mountain of words was coming and that we would lose any doubt that the task was, in fact, impossible.

With arrival of the fall semester, a post-doc who had previously worked with the Center for Studies in Higher Education dropped by the office. She was back in the area, spending most of her time with her toddler son but was interested in working part time while she wrote grant proposals that would provide ongoing support. In our conversation, I spoke with her about the text analysis challenge that kept me paralyzed. She mentioned that most graduate students in her doctoral program were using Atlas.ti and were raving about its qualities. Internet surfing confirmed that it was considered to be one of two really good options. Sure it was expensive, but the ability to create “families” of responses from groups sharing characteristics and to write “super codes” that would automatically code much of the file, made it a very attractive product.

The software was ordered, received, installed and training began. After a couple days self-administered training, learning terms like hermeneutic unit (pronounced 'Her`me*neu"tics'), it dawned on me that I had not seen the equivalent of an ID used in a training file. Surely that was only because they were training files. I had dreamed that I could extract all comments that had been made in response to a question, develop codes and process those responses, then return the resulting codes to the data file where they could be linked to any and all of the hundreds of variables available. I had to be able to attach an ID to the codes to make that work. So, how would I attach an ID using this program? It couldn't be hard to do because the need was so obvious, right? Wrong. After days of forum exchanges, web searches, phone calls, and meetings with other users, the inconvenient truth became clear. Record identification could be done using document name, but I would have to form about 45,000 primary documents, name them, open them one-by-one, and code each. When finished, I could return the 45,000 individual files to the survey data file as a variable. The project was a failure on which I had wasted considerable time and money. And worse, the analysis was the subject of my conference presentation to valued peers. It too would be a failure.

Act 3: If You Only Have a Hammer, Then Every Problem Looks Like a Nail

Could I possibly turn these messy words and phrases into clean orderly numbers and learn something from those numbers? Having no idea what else I might try, I thought what the heck. Let's look at these comments as quantitative values. The comments have at least two numerically measurable characteristics. First, the comment either exists or it does not. Second, if it exists, it has length and is relatively longer or shorter than other comments. If the Bush administration's support of search and wiretapping without a judge's approval can be called “warrantless” than the examination of length of student comments while ignoring the meaning of those comments might be called “meaningless” analysis.

True to what I had been trained, I generated a few reasonable sounding questions. Was the likelihood that a student would comment related to student characteristics: campus of attendance, questionnaire module version received, class level, matriculation status, year in school, sex, disciplinary field and race/ethnicity? And, if they did offer a comment, was the length of that comment associated with those variables? Whether or not relationships existed, the analysis was one that could be done and the results could be shared to fill time during a 40 minute conference paper session. Just between you and me, analyses like these also happen to be the sort of thing I geekout on.

<Insert Table 1 About Here>

Table 1 displays, by column, enrollment, number of responding students and whether students made one, two or three comments. The patterns in Table 1 are better shown graphically. Even a quick glance at Figure 1 shows clearly that the likelihood that a student at a campus will make a comment was related to the campus response rate. The lines from survey participation to first comment, don't cross. The next fact that is clear is the subsequent probability that a student will make a second and then a third comment is consistent across campuses. The retention patterns, or loss patterns if you're the half-empty type, are remarkably parallel. The most remarkable patterns are those for Berkeley and Irvine campuses. They so completely overlap that only one line is visible. If there is a message to take away, it is that about 78% of those who participated in the survey made a comment and that about 88% of those students made a second, and about 86% of those made a third. The retention/attrition pattern was very consistent. I don't know what that means, but I find it interesting. If nothing else, it can be used to estimate the number of comments that will be made.

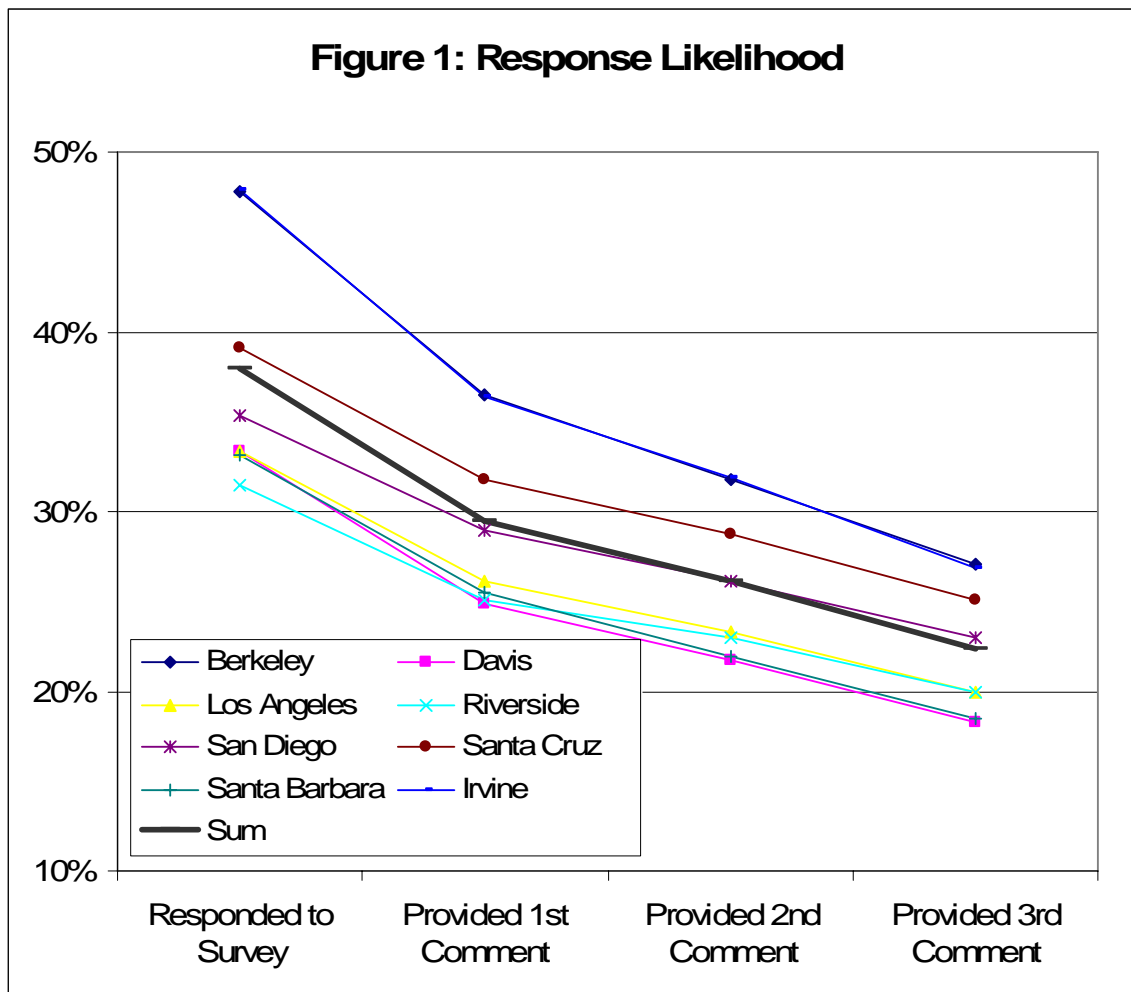
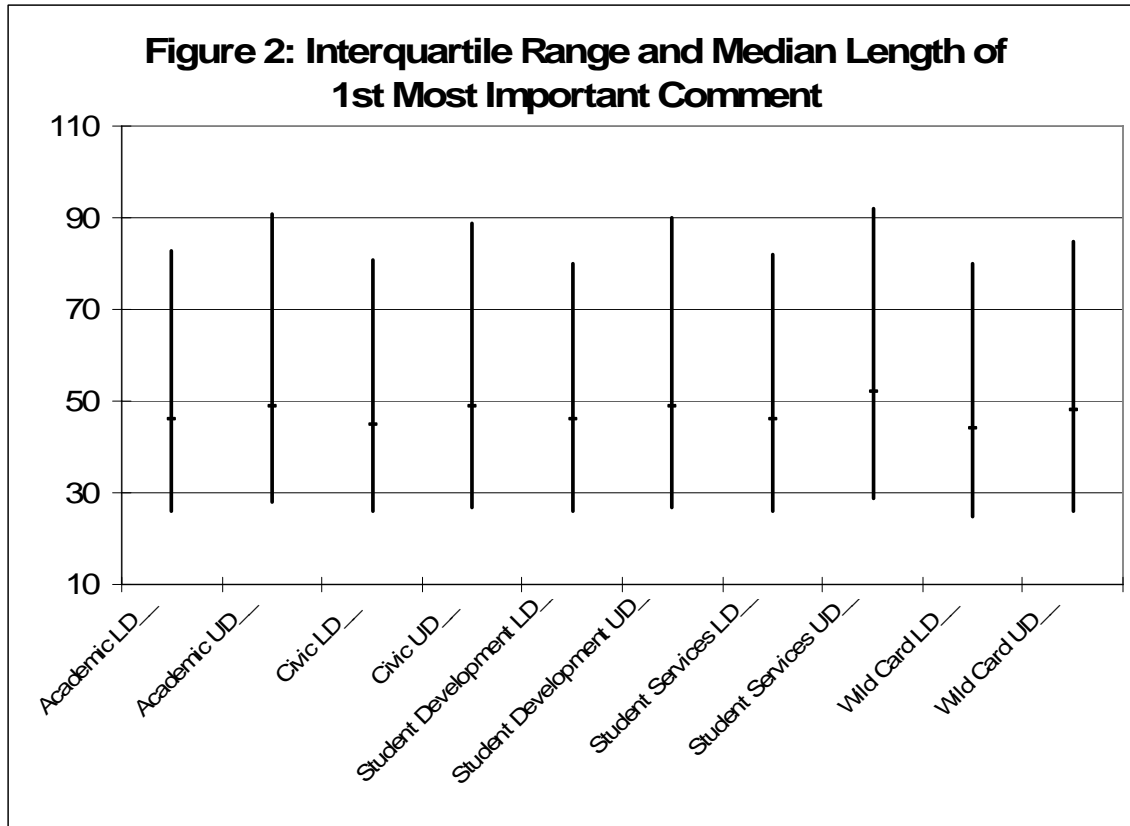


Table 2 pushes this analysis even further toward the obscure, but interesting in that factoid way that occupies too much of our news and brain cells. Students at Davis, Irvine, Berkeley and Santa Barbara were less likely to offer comments even though Berkeley and Irvine students were the most likely to respond to the survey. Interestingly, students were more likely to comment when the questionnaire was shorter. That is

interesting because they would have had little information about questionnaire length when they came to that item. Lower- and upper-division students were equally likely to comment but upper-division students were more likely to make only one or two comments. There were no significant differences by matriculation status—transfer and native students were very similar. And year in attendance was also unrelated to likelihood to comment or offer multiple comments. Males were more likely to comment and to offer three comments. Mathematics and computer science students were the least likely to comment and especially least likely to make three comments. Humanities, social sciences and area and ethnic studies majors were more likely to comment and to make three comments. Differences by race and ethnicity sorted African American, Pacific Islander and Japanese students as the most likely to comment and Thai, Chinese and Vietnamese as the least likely. None of these differences explained more than 6% of variation in the patterns observed.

<Insert Tables 2 And 3 About Here>

If you haven't yet reached your limit, Table 3 repeats the analysis for length of comment made and by doing so makes a remarkable discovery. The best predictor of the length of a third comment is the length of the second and, in turn, the first. Whatever length of comment a group of students made, tended to be consistent from first to third comments. OK, maybe not so remarkable at a global level, but what is remarkable is that this consistency was maintained when there were group differences. For example, UC Davis students made longer first comments than Riverside or San Diego students by almost 10 characters or two words and they made similarly longer second and third comments. Maybe a reason for campus differences comes to your mind. It doesn't come to mine and, even if it did, I wouldn't understand why the difference should extend to second and third comments. I imagine that I can see why some of the other patterns were found—some, but certainly not all. Upper-division and senior students made longer comments. Freshmen matriculates and transfers did not differ, they seldom do. Females made slightly longer comments than males, and business administration majors made shorter comments than all other majors. Area and ethnic studies and agriculture and architecture students made longer comments. When examining length of comments by race/ethnicity using the more crude federal categories, shorter comments were made by Asians/Pacific Islanders and Hispanics and longer comments were made by American Indians/Alaskan Natives and Whites/Caucasians. When seen in greater racial/ethnic detail, the shortest comments were by Chinese, Korean and Vietnamese students. Overall, the largest difference was by student level. Upper-division students made longer comments and the difference was very consistent across questionnaire versions (Figure 2). Again, differences observed for first comments extended to second and third comments.



See how desperate I was to find something to present? In my defense, there is at least one useful application for this esoteric, quantitative work. The length of a comment permitted on the questionnaire form was not limited. Students could write on and on and many did. One way to help control the analytical burden in the future is to limit comment length but that should be done sparingly and should have minimal impact. If we use a 90th percentile standard, a text box permitting about 150 characters, or about 25 words, should be sufficient. Don't give up. There is another act.

Act 4: Skinning Cats

Sometimes, we can't see alternative solutions because we are too close, too heavily invested in one strategy. For this project, I was too heavily invested in the idea that we should code every student's remarks. I thought that it was important that we showed our respect for the students by doing so. It was, in fact, the least that we could do, but the doing was impossible. If only there were a way to carefully consider relatively few remarks and then be able to speak with confidence about what all had said. Maybe, focus on a few randomly selected ... Oh yeah, Stat 101: Random sampling, population frames, generalization. Why not process random samples taken to address specific needs? For example, was it really necessary to ask for the three most important things we might do to improve undergraduate education? Do we learn anything different or only something more by asking for three?

That answer could be found simply enough. We could randomly select the first remarks of 100 students and the third remarks of a different 100 students and compare them. Coding would be made easier because we could use Monica Lopez's topology—no

need to reinvent that wheel. Could it really be that easy? I hoped so because we needed to decide quickly whether we could save students' questionnaire completion time by asking for one recommendation instead of three. The random selection of first and third remarks was made using SAS Proc Surveyselect and specifying that each of the two samples would total 100 with sample units reflecting the enrollment distribution of the campuses: 15 from Berkeley, 14 from Davis, etc. The sampling frame would also be limited to students making three entries. (Good news. The comments as quantitative values analysis had a useful application.)

<Insert Table 4 About Here>

In less than one day, I prepared the distributions shown as Table 4. As you can see, the distribution of comments for the first and third most important recommendations were virtually identical. Based on this relatively easy analysis, we don't need to ask for three recommendations. At least at the level of coded results, we don't learn anything new. The results were redundant. But did Monica's coding structure fit the 2006 group? It looked like a reasonable and fairly comprehensive coding structure, but it had been developed in response to a different question and it was three years old, forever in undergraduate years.

It isn't a simple task to compare the 2006 results with those from 2003 because it was common for 2003 students to offer multiple recommendations and those recommendations crossed category lines. Monica Lopez's computations were also not crystal clear and at least one significant error was found. (So she wrote 1% when she meant to write 10%. Who among us hasn't done as much?) That aside, my best guesses where comparison was possible showed similar results. Advising was listed by 9% of 2006 students and 13% of 2003 students. Remarks about classes were 19% of 2006 and 23% of 2003. Faculty were the subject of 8% of 2006 and 13% of 2003 comments. Instruction was 6% of 2006 and 7% of 2003. I suspect that Monica and I differed in placing comments about faculty instruction in faculty or instruction categories. Nevertheless, the profiles were similar and more importantly, the Lopez categories fit 2006 very well. The challenge of comparing results did show one way in which a question asking for three comments was preferable to a question asking for one. When asked for three items, students were not so inclined to list multiple topics in one response. Don't you hate it when they do that?

Maybe the enumerative analysis of the third act could show useful directions for this technique. In particular, there was a clear tendency for lower-division students to use fewer characters when making remarks. They were equally likely to make comments but those comments were shorter. Did the meaning of those comments differ? Using small random samples, it is a question that can be addressed and reported in less than one day. A random sample of 100 was taken from each group, again using campus proportions, but this time not limiting analysis to those making three comments. I did limit this analysis to student who matriculated from high school, not transfers. A few hours later, the results were summarized in Table 5. The level-three coding detail showed that:

- Freshmen were more likely to talk about access to courses; academic support, tutoring, and assistance; housing, on-campus food; and campus climate, especially the need to increase diversity.

- Seniors were more likely to talk about the need for smaller classes; instruction—curriculum, class content; amount, management, and delivery of services; and facilities (need for social centers, too much construction)

At a higher level of clustering, level 2, fewer differences were observed. Seniors were more likely to comment about instruction, delivery of services, and nonacademic facilities.

<Insert Table 5 About Here>

This very efficient strategy that can easily be applied to dozens of problems. Do transfer students in their first year and first-year freshmen differ? What about men and women? Do underrepresented minority students offer different advice than majority students? How about STEM students? Their experience differs in many ways. Are these differences reflected in their advice for improving the university experience? Frankly, using random samples and the existing coding structure means that all these questions could be answered in less than one week.

But what about those tens of thousands of student comments that would not be selected as part of a random sample? Aren't they as valuable? Don't we owe students the respect of reading what they wrote? That question reminds me of a discovery I made about myself during this project. I am not a fan of qualitative analysis generally and less of a fan of qualitative text analysis of student comments on questionnaires. They are awkward, messy things that were even illegible and had to be retyped in the olden days when I became an institutional researcher. That I was predisposed against open-ended items was not the discovery. I knew that, but as I was struggling to place well worded and insightful bits of student wisdom into categories identified by a few key words, I realized that the activity was almost exactly the opposite of what I had been trying to do for most of my professional life. Not because the material was text but because the task had a different goal.

Most of my work is about trying to increase the value of enumerative data by finding patterns, context, relationships, or generally translating them into something of greater meaning to support a decision or to make a useful contribution to our understanding. The task of assigning codes to more fully expressed thoughts is one of removing value by changing something full and rich into broadly stated clusters labeled by a few words or a phrase. It seems so inadequate. Here are a few examples.

Before coding, the comment was, "Have the Paris trip be accessible to more students. If the same few students are already guaranteed to go every year, it is not fair to the other 150 students who will never have the opportunity." After coding, this specific recommendation is "Resources, Academic--Enrichment, Opportunities for." In a similar way, the very specific and probably useful suggestion, "Make a bike path through the Pavilion area between Girvetz Hall and the Davidson Library" becomes "Transportation, University Provided--Amount and Type." Or, "Light a fire under the arses of professors so they at least pretend like teaching is ONE of their priorities" does nothing more than increase the frequency count of "Quality of Instruction." There is so much lost in this process. It is true that we can offer these and a few other comments as examples for the category to try and show that the cluster is actually full and rich, but does that really work? There is an alternative. The full meaning doesn't have to be lost.

It is the practice of some offices, SARI at UC Davis for example, to make all comments available to those responsible for programs. For example, comments from majors and alumni go to department chairpersons. We could do the same with UCUES results. The actual comments could be reported at levels of aggregation where they would be most effective. For example, it would be easy enough to read all the comments by female Vietnamese students born outside the States and their comments might be uniquely helpful to an outreach advisor. Likewise, the comments of History majors at Riverside can, and probably should, be shared with the committee reviewing that academic program. In cases like these, it isn't necessary or particularly useful to report distributions of comments by category to faculty, administrators, and staff. Likewise, random samples of unmolested comments could be given upon approved request to meet the needs and interests of researchers. The dual strategy of coding randomly selected comments to meet specific needs and to share the full text of comments with interested parties, seems to be a useful compromise. It is also something that can, in fact, be accomplished.

So, this brings me to the end of this paper. Like most institutional research projects it ended someplace different than was planned. But this one ended further from the planned goal than most. Along the way, I grew to know and respect the work of a graduate student about who I really know nothing more than a name. I waded from the shore into the ocean of qualitative data analysis intending to ride the waves on my fancy software surfboard, and I very nearly drowned. I fell back onto the simple skills I had and dog paddled. But in the end. I found that there was at least one alternative and that alternative worked for me.

Table 1: Likelihood of Responding, Commenting, and Making Multiple Comments

Campus	Enrollment	Responded to Survey	Provided 1st Comment	Provided 2nd Comment	Provided 3rd Comment
Frequencies					
Berkeley	22,424	10,717	8,186	7,129	6,068
Davis	21,814	7,270	5,419	4,733	3,993
Los Angeles	23,644	7,882	6,188	5,507	4,711
Riverside	13,888	4,368	3,488	3,185	2,777
San Diego	20,238	7,159	5,849	5,293	4,659
Santa Cruz	13,112	5,129	4,173	3,766	3,287
Santa Barbara	17,513	5,800	4,458	3,847	3,245
Irvine	19,680	9,430	7,168	6,274	5,289
Sum	152,313	57,755	44,929	39,734	34,029
Percent of All Students					
Berkeley		48%	37%	32%	27%
Davis		33%	25%	22%	18%
Los Angeles		33%	26%	23%	20%
Riverside		31%	25%	23%	20%
San Diego		35%	29%	26%	23%
Santa Cruz		39%	32%	29%	25%
Santa Barbara		33%	25%	22%	19%
Irvine		48%	36%	32%	27%
Sum		38%	29%	26%	22%
Retention					
Berkeley		48%	76%	87%	85%
Davis		33%	75%	87%	84%
Los Angeles		33%	79%	89%	86%
Riverside		31%	80%	91%	87%
San Diego		35%	82%	90%	88%
Santa Cruz		39%	81%	90%	87%
Santa Barbara		33%	77%	86%	84%
Irvine		48%	76%	88%	84%
Sum		38%	78%	88%	86%

Table 2: Likelihood of Making 1, 2 or 3 Comments

	Cramer's V	Only One Comment	Two Comments	Three Comments	Provided a Comment	Did Not Provide Comment
Campus						
San Diego		8%	9%	65%	82%	18%
Santa Cruz		8%	10%	64%	81%	19%
Riverside		7%	10%	63%	80%	20%
Los Angeles		9%	11%	59%	78%	22%
Santa Barbara		11%	11%	55%	77%	23%
Berkeley		10%	10%	56%	76%	24%
Irvine		10%	11%	55%	76%	24%
Davis		10%	11%	54%	74%	26%
Total	0.0474	9%	10%	58%	78%	22%
Module						
Wild Card LD		8%	9%	63%	80%	20%
Student Development LD		8%	10%	61%	79%	21%
Academic UD		10%	12%	56%	78%	22%
Academic LD		8%	9%	62%	78%	22%
Wild Card UD		10%	11%	56%	78%	22%
Civic UD		10%	12%	56%	78%	22%
Student Development UD		10%	11%	56%	77%	23%
Civic LD		8%	9%	60%	77%	23%
Student Services UD		11%	11%	55%	76%	24%
Student Services LD		9%	9%	58%	76%	24%
Total	0.0383	9%	10%	58%	78%	22%
Class Level Division						
Lower Division		8%	9%	61%	78%	22%
Upper Division		10%	11%	56%	77%	23%
Total	0.059	9%	10%	58%	78%	22%
Matriculation Status						
Freshman		9%	10%	59%	78%	22%
Transfer		11%	11%	53%	76%	24%
Total	0.036	9%	10%	58%	78%	22%
Campus Level						
Second		8%	9%	62%	79%	21%
Fourth		10%	11%	58%	78%	22%
Third		10%	11%	56%	77%	23%
First		9%	9%	58%	76%	24%
Total	0.031	9%	10%	58%	78%	22%
Sex						
Male		9%	9%	60%	78%	22%
Female		9%	11%	56%	77%	23%
Total	0.0457	9%	10%	58%	78%	22%
Disciplinary Field (UD)						
Area Ethnic		10%	11%	61%	82%	18%
Humanities/Soc Sci		10%	12%	57%	79%	21%
Business Adm		11%	11%	54%	77%	23%
Bio and Physical Sci		10%	12%	54%	76%	24%
Agric/Architecture		9%	9%	57%	75%	25%
Engineering		10%	9%	54%	74%	26%
Math/Computer Sci		11%	11%	49%	72%	28%
Total	0.0371	10%	11%	56%	77%	23%

Table 2: Likelihood of Making 1, 2 or 3 Comments

	Cramer's V	Only One Comment	Two Comments	Three Comments	Provided a Comment	Did Not Provide Comment
Race/Ethnicity						
Amer Indian/Alaskan		6%	8%	66%	80%	20%
Black/African American		8%	11%	61%	80%	20%
Pacific Islander		7%	9%	63%	80%	20%
Japanese		9%	10%	61%	80%	20%
White/Caucasian		10%	11%	58%	79%	21%
Other		10%	11%	58%	79%	21%
Korean		9%	10%	59%	78%	22%
Decline		9%	10%	60%	78%	22%
East Indian/Pakistani		8%	10%	60%	78%	22%
Pilipino		8%	9%	60%	77%	23%
Chicano/Mexican America		10%	10%	57%	77%	23%
Latino		10%	11%	56%	77%	23%
Thai/Other Asian		9%	9%	58%	76%	24%
Chinese		9%	10%	56%	75%	25%
Vietnamese		8%	10%	56%	74%	26%
Total	0.0332	9%	10%	58%	78%	22%
IPEDS Race/Ethnicity						
Amer Indian/Alaskan		6%	8%	66%	80%	20%
Black/African American		8%	11%	61%	80%	20%
White/Caucasian		10%	11%	58%	79%	21%
Other		10%	11%	58%	79%	21%
Decline		9%	10%	60%	78%	22%
Hispanic		10%	11%	57%	77%	23%
Asian/Pacific Islander		8%	10%	58%	76%	24%
Total	0.0275	9%	10%	58%	78%	22%

Table 3: Length of Comments in Character Count

	1st				Median Words*	2nd				3rd			
	25th	75th	50th	90th		25th	75th	50th	90th	25th	75th	50th	90th
Overall	27	85	48	151	8.0	28	84	48	150	26	83	45	155
Campus													
Davis	30	96	54	166	9.0	31	94	52	160	30	94	51	170
Santa Barbara	29	88	50	153	8.3	31	88	52	151	28	87	49	150
Santa Cruz	28	88	50	156	8.3	29	92	51	163	27	92	49	179
Berkeley	26	83	46	147	7.7	28	82	48	145	26	81	45	148
Irvine	25	84	46	152	7.7	26	81	45	146	24	78	42	146
Los Angeles	25	82	46	140	7.7	28	82	47	143	26	80	45	143
San Diego	25	82	45	150	7.5	27	81	46	148	24	79	42	155
Riverside	25	79	44	149	7.3	27	81	46	142	25	79	43	153
Module													
Academic LD__	26	83	46	142	7.7	27	78	46	138	25	76	42	139
Academic UD__	28	91	49	158	8.2	30	92	51	164	28	92	49	174
Civic LD__	26	81	45	136	7.5	27	79	45	137	25	75	43	137
Civic UD__	27	89	49	164	8.2	30	92	52	162	29	94	51	179
Student Development LD_	26	80	46	141	7.7	27	78	45	133	24	76	42	138
Student Development UD_	27	90	49	164	8.2	29	89	49	157	28	90	49	169
Student Services LD__	26	82	46	141	7.7	27	79	46	136	25	76	42	142
Student Services UD__	29	92	52	164	8.7	30	92	52	166	28	93	50	172
Wild Card LD__	25	80	44	133	7.3	27	76	44	136	24	72	41	131
Wild Card UD__	26	85	48	158	8.0	30	87	50	152	27	87	47	158
Class Level Division													
Lower Division	26	81	46	140	7.7	27	78	45	136	24	75	42	137
Upper Division	27	89	49	161	8.2	30	90	51	161	28	92	49	171
Matriculation Status													
Freshman	26	85	47	148	7.8	28	83	48	147	26	81	45	151
Transfer	27	86	48	164	8.0	29	89	49	161	27	91	48	172
Campus Level													
First	25	76	44	130	7.3	26	74	43	125	23	69	40	123
Second	26	84	46	142	7.7	27	80	46	138	25	77	43	142
Third	27	85	48	153	8.0	29	86	49	155	27	86	46	161
Fourth	27	92	50	165	8.3	30	92	52	164	28	93	50	174
Sex													
Female	28	87	49	153	8.2	30	87	50	153	27	86	47	159
Male	25	82	46	148	7.7	26	80	46	145	25	79	43	148
Disciplinary Field (UD)													
Agric/Architecture	31	103	55	172	9.2	34	97	54	170	31	103	58	191
Area Ethnic	29	97	55	171	9.2	29	99	51	172	32	95	55	181
Math/Computer Sci	28	91	51	163	8.5	29	91	51	169	26	91	46	174
Humanities/Soc Sci	28	91	50	163	8.3	31	92	52	163	28	93	50	174
Bio and Physical Sci	25	87	48	154	8.0	29	90	50	156	27	89	48	169
Engineering	27	86	47	158	7.8	28	85	49	155	26	86	48	157
Business Adm	24	73	41	141	6.8	26	75	45	129	26	78	43	145
Race													
Amer Indian/Alaskan	28	97	54	172	9.0	26	96	49	178	25	96	49	175
White/Caucasian	30	95	53	169	8.8	31	94	53	165	29	95	51	174
Pacific Islander	29	88	52	166	8.7	29	87	48	145	27	90	48	164
Decline to State	29	92	51	163	8.5	30	90	51	165	27	91	49	166
Other	29	94	50	171	8.3	31	91	50	164	28	92	48	166
Latino	27	86	49	150	8.2	29	91	51	160	26	84	45	159
East Indian/Pakistani	26	88	48	151	8.0	28	83	49	149	26	81	45	156
Black/African American	27	82	47	144	7.8	29	86	48	154	27	81	48	166
Japanese	27	81	47	143	7.8	29	81	48	146	26	75	41	145
Pilipino	27	83	46	149	7.7	27	81	46	139	26	79	43	143

Table 3: Length of Comments in Character Count

	1st		50th	90th	Median Words*	2nd		50th	90th	3rd		50th	90th
	25th	75th				25th	75th			25th	75th		
Thai/Other Asian+C26	26	83	45	146	7.5	27	79	45	137	23	77	43	142
Chicano/Mexican Amer	25	78	44	137	7.3	27	77	46	132	25	76	43	139
Chinese	23	73	41	131	6.8	24	72	42	126	22	69	38	129
Korean	23	73	41	129	6.8	25	71	41	126	23	69	39	132
Vietnamese	22	75	41	134	6.8	25	74	43	131	23	74	41	133
IPEDS Race/Ethnicity													
Amer Indian/Alaskan	28	97	54	172	9.0	26	96	49	178	25	96	49	175
White/Caucasian	30	95	53	169	8.8	31	94	53	165	29	95	51	174
Decline to State	29	92	51	163	8.5	30	90	51	165	27	91	49	166
Other	29	94	50	171	8.3	31	91	50	164	28	92	48	166
Black/African American	27	82	47	144	7.8	29	86	48	154	27	81	48	166
Hispanic	26	79	46	139	7.7	28	81	47	138	25	78	44	145
Asian	24	77	43	136	7.2	26	75	44	132	24	73	41	136

* At an average length of length of 5 characters and 1 space.

Correlations	1st &	1st &	2nd &
	2nd	3rd	3rd
Campus	0.90	0.93	0.98
Module	0.93	0.90	0.96
Disciplinary Field (UD)	0.91	0.87	0.81
Race	0.89	0.91	0.86
IPEDS Race/Ethnicity	0.85	0.90	0.94

Table 4: Comparison of Coded* Suggestions Identified by Respondents as First Most Important and Third Most Important

<u>First Most Important</u>				<u>3rd Most Important</u>				<u>Difference</u>
<u>L1</u>	<u>L2</u>	<u>L3</u>	<u>#</u>	<u>L1</u>	<u>L2</u>	<u>L3</u>	<u>#</u>	
Level 1								
a			69	a			68	1
s			19	s			18	1
u			8	u			8	0
z			3	z			5	-2
o			1	o			1	0
Levels 1 and 2								
a	a		9	a	a		9	0
a	c		18	a	c		18	0
a	f		7	a	f		8	-1
a	i		6	a	i		6	0
a	l		6	a	l		6	0
a	o		1	a	o		1	0
a	r		2	a	r		2	0
a	s		18	a	s		18	0
a	t		1	a	t		1	0
o	o		1	o	o		1	0
s	a		1	s	a		1	0
s	d		1	s	d		1	0
s	f		1	s	f		1	0
s	h		9	s	h		9	0
s	p		5	s	p		6	-1
s	t		1	s	t		1	0
u	e		1	u	e		1	0
u	f		1	u	f		1	0
u	l		6	u	l		6	0
z	z		5	z	z		3	2
Levels 1,2, and 3								
a	a	g	1	a	a	g	1	0
a	a	k	2	a	a	k	2	0
a	a	p	2	a	a	p	2	0
a	a	q	4	a	a	q	4	0
a	c	a	10	a	c	a	10	0
a	c	k	2	a	c	k	2	0
a	c	s	6	a	c	s	6	0
a	f	c	1	a	f	c	1	0
a	f	i	3	a	f	i	2	1
a	f	m	2	a	f	m	2	0

Table 4: Comparison of Coded* Suggestions Identified by Respondents as First Most Important and Third Most Important

First Most Important				3rd Most Important				Difference
L1	L2	L3	#	L1	L2	L3	#	
a	f	n	1	a	f	n	1	0
a	f	s	1	a	f	s	1	0
a	i	c	4	a	i	c	4	0
a	i	q	2	a	i	q	2	0
a	l	c	1	a	l	c	1	0
a	l	d	3	a	l	d	3	0
a	l	u	2	a	l	u	2	0
a	o	g	1	a	o	g	1	0
a	r	r	2	a	r	r	2	0
a	s	a	6	a	s	a	6	0
a	s	f	3	a	s	f	3	0
a	s	g	2	a	s	g	2	0
a	s	m	2	a	s	m	2	0
a	s	n	4	a	s	n	4	0
a	s	s	1	a	s	s	1	0
a	t	m	1	a	t	m	1	0
o	o	o	1	o	o	o	1	0
s	a	o	1	s	a	o	1	0
s	d	d	1	s	d	d	1	0
s	f	p	1	s	f	p	1	0
s	h	a	1	s	h	a	1	0
s	h	c	1	s	h	c	1	0
s	h	f	5	s	h	f	4	1
				s	h	h	1	-1
s	h	m	1	s	h	m	1	0
s	h	s	1	s	h	s	1	0
s	p	g	6	s	p	g	5	1
s	t	m	1	s	t	m	1	0
u	e	g	1	u	e	g	1	0
u	f	g	1	u	f	g	1	0
u	l	n	4	u	l	n	4	0
u	l	p	1	u	l	p	1	0
u	l	s	1	u	l	s	1	0
z	z	z	3	z	z	z	5	-2

Based on responses by 200 randomly selected students. The replies for most important were analyzed for a sample of 100 and the

Note: replies for the third most important were analyzed for a different

* See Appendix for description of coding.

Table 5: Comparison of Coded* Suggestions Identified by 1st Year and 4th Year High School Matriculants

First Year				Fourth Year				Difference	Prob.>
L1	L2	L3	#	L1	L2	L3	#		
Level 1									
a			63	a			68	-5	0.72
s			11	s			11	0	
u			22	u			18	4	0.70
z			4	z			3	1	0.50
o				o					
Levels 1 and 2									
a	a		14	a	a		9	5	0.80
a	c		16	a	c		19	-3	0.65
a	f		7	a	f		7	0	
a	i		3	a	i		12	-9	0.98
a	l		1	a	l		4	-3	0.82
a	o		2	a	o		1	1	0.50
a	r		2	a	r		5	-3	0.78
a	s		14	a	s		10	4	0.64
a	t		4	a	t		1	3	0.69
s	a		1	s	a		0	1	0.50
s	d		0	s	d		3	-3	0.88
s	f		1	s	f		1	0	
s	h		7	s	h		7	0	
s	l		1	s	l		0	1	0.50
s	n		0	s	n		1	-1	0.50
s	p		1	s	p		3	-2	0.69
s	t		0	s	t		2	-2	0.76
u	a		0	u	a		3	-3	0.88
u	e		1	u	e		0	1	0.50
u	f		5	u	f		3	2	0.64
u	l		16	u	l		12	4	0.73
z	z		4	z	z		3	1	0.50
Levels 1,2, and 3									
a	a	k	4	a	a	k	1	3	0.82
a	a	q	9	a	a	q	7	2	0.60
a	a	t	1	a	a	t	1	0	
a	c	a	9	a	c	a	10	-1	0.50
a	c	k	4	a	c	k	0	4	0.94
a	c	s	3	a	c	s	9	-6	0.93
a	f	c	0	a	f	c	1	-1	0.50
a	f	i	2	a	f	i	4	-2	0.66
a	f	m	4	a	f	m	2	2	0.66
a	f	n	1	a	f	n	0	1	0.50
a	i	c	1	a	i	c	10	-9	0.99

Table 5: Comparison of Coded* Suggestions Identified by 1st Year and 4th Year High School Matriculants

First Year				Fourth Year				Difference	Prob.>
L1	L2	L3	#	L1	L2	L3	#		
a	i	m	0	a	i	m	1	-1	0.50
a	i	q	2	a	i	q	0	2	
a	i	r	0	a	i	r	1	-1	0.50
a	l	c	0	a	l	c	1	-1	0.50
a	l	d	1	a	l	d	3	-2	0.69
a	o	g	2	a	o	g	1	1	0.50
a	r	d	0	a	r	d	2	-2	0.76
a	r	g	0	a	r	g	1	-1	0.50
a	r	r	2	a	r	r	2	0	
a	s	a	2	a	s	a	1	1	0.50
a	s	c	2	a	s	c	0	2	0.76
a	s	f	0	a	s	f	1	-1	0.50
a	s	g	0	a	s	g	1	-1	0.50
a	s	n	3	a	s	n	6	-3	0.75
a	s	s	7	a	s	s	1	6	0.96
a	t	m	4	a	t	m	1	3	0.82
s	a	o	1	s	a	o	0	1	0.50
s	d	d	0	s	d	d	3	-3	0.88
s	f	a	1	s	f	p	1	0	
s	h	a	1	s	h	a	1	0	
s	h	d	1	s	h	c	0	1	0.50
s	h	f	3	s	h	f	0	3	0.88
s	h	g	1	s	h	g	0	1	0.50
s	h	m	1	s	h	m	0	1	0.50
s	l	g	1	s	l	g	0	1	0.50
s	n	a	0	s	n	a	1	-1	0.50
s	p	g	1	s	p	g	3	-2	0.69
s	t	d	0	s	t	d	2	-2	0.76
u	a	g	0	u	a	g	3	-3	0.88
u	e	g	1	u	e	g	0	1	0.50
u	f	g	5	u	f	g	3	2	0.64
u	l	c	0	u	l	c	2	-2	0.76
u	l	d	3	u	l	d	0	3	0.88
u	l	g	0	u	l	g	1	-1	0.50
u	l	p	13	u	l	p	9	4	0.75
z	z	z	4	z	z	z	3	1	0.50

Note: Based on responses by 200 randomly selected students. The replies for most important were analyzed for a sample of 100 and the replies for the third most important were analyzed for a different

* See Appendix for description of coding.

Appendix 1: Coding System for Open-Ended Responses

	Level 1	Level 2	Level 3
A O O	Academic		
A A G	Advising		
A A K		Access, Availability & Amount	
A A P		Provider	
A A Q		Quality and Content, Transition Courses	
A A T		Tracking, Monitor Progress	
A C G	Classes		
A C K		Access/Location	
A C A		Availability	
A C S		Class Size	
A F G	Faculty		
A F C		Caring -- Concern for students, teaching	
A F I		Interaction -- Ability, opportunity and willingness to interact	
A F M		Competence, teaching	
A F N		Input -- Collecting and using student judgment	
A F O		Opinion -- Don't preach political agenda	
A F S		Staffing Mix -- More professors, fewer lecturers	
A I G	Instruction		
A I C		Curriculum/Class Content	
A I M		Staffing Mix to Deliver	
A I Q		Quality of Instruction	
A I R		Grading	
A L G	Climate		
A L C		Competitive, Too Much	
A L D		Diversity, Need More Admissions/Emphasis/Requirements	
A L L		Diversity, Need Less Admissions/Emphasis/Requirements	
A L U		Uncaring / Assembly Line	
A O G	Other		
A R G	Requirements		
A R A		Admission to Major	
A R B		Bureaucracy, Paperwork	
A R D		Gen Ed, Too Much	
A R R		Rigidity	
A S G	Resources, Academic		
A S A		Facilities: Lab, Class, Study	
A S C		Computer, Wireless Networks	
A S F		Funding for Academic Department(s)	
A S M		Majors, Offer More/Different/Change	
A S N		Enrichment, Opportunities for Academic	
A S S		Support, Academic (Tutoring/Assistance)	
A T G	Teaching Assistants		
A T C		Coordination with Faculty	
A T L		Language Proficiency	
A T M		Competency	
S O O	Amount, Administration & Management of Services		
S A O	Other		
S D D	Delivery of Services, Other/General		

Appendix 1: Coding System for Open-Ended Responses

Level 1			
		Level 2	
		Level 3	
S	F	G	Financial Aid
S	F	A	Amount & Type
S	F	D	Delivery of Services
S	F	P	Processes & Procedures
S	H	G	Housing and On-Campus Food Services
S	H	A	Availability
S	H	C	Condition
S	H	D	Delivery of Services
S	H	F	Food
S	H	M	Amount & Type
S	H	S	Cost
S	L	G	Health Services, Mental Health
S	N	A	Information Availability
S	P	G	Parking
S	T	G	Transportation, University Provided
S	T	D	Delivery of Services
S	T	M	Amount & Type
U	C	G	General Campus
U	C	O	Other
U	E	G	Employment Opportunities
U	A	G	Facilities (Social Centers, Outside Commercial, Amount Construction, Etc.)
U	F	G	Fees, Required & Cost of Attendance, How Spent
U	H	G	Housing Off Campus
U	L	G	Campus Climate (Nonacademic)
U	L	C	Student Centered, Students Treated as People
U	L	D	Diversity, Increase (Nonacademic/Not Admissions) and Tolerance
U	L	E	Campus Size / Enrollment
U	L	F	Safety
U	L	L	Campus Location (Affordability, Off-Campus Entertainment)
U	L	N	Enrichment, Opportunities for Academic and Related (e.g., leadership, service)
U	L	P	Political, Entertainment and Social Climate, Quality of
U	L	S	Staff/Student Relationships
U	T	G	Transportation, Not University+F69
O	O	O	Other

* Based on the work of Monica Lopez (May, 2005) *Advice from Students on Improving Their Undergraduate Experience at the University of California*