Initial Evaluation of Cabrillo College Spring 2005 Call Center

(Draft)

Jing Luan
Cathy Summa
Marcy Wieland

EXECUTIVE SUMMARY

Cabrillo College conducted outreach calls through a call center staffed by 50 volunteers to thousands of applicants who applied for admissions but had not registered one month prior to the start of spring 2005 semester. This outreach effort was in response to an anticipated decrease in enrollment due to fee increase. Following the calls, the college evaluated the Call Center activities. A total of four (4) research questions were identified to address yield, productivity, process improvement and the use of data mining to pre-classify future applicants into groups of likely and unlikely registrants.

In responding to the four (4) research questions, the study found a statistically significant 5% higher registration rate in a four way post-facto control and experimental analysis (page 4). However, this significance should also take into consideration a few college initiated changes, although no specific impact to any of the four way analyses was immediately perceivable. For example, the college implemented online application in spring 2005, which increased the total applicants by 12% as compared to the spring a year earlier.

Through examining the applicants by the manner in which they were reached by the Call Center, the study found a noticeable difference in registration rate that was attributable to the presence of the Call Center (page 5). Those who spoke directly with the Call Center staff had at least a 10% increase in registering for classes compared to those who had no contact with the Call Center. Yet in answers to Question Three on Productivity, the increase was found, at best, to have helped maintain the level of FTES, a productivity measure (page 7). However, it can be viewed that without the Call Center, there could have been a perceivable drop in both headcounts and FTES in spring 2005. From a summative evaluation perspective, the Call Center receives a B+.

After removing the applicants who interacted with the Call Center directly, data mining analysis identified six (6) rules for applicants who would register for classes and another six (6) for those who would not (Appendix B). Further, the data mining analysis built a predictive model with an 85% accuracy to predict those who would be less likely to register in a future semester (page 16).

The Call Center was a success in mobilizing an army of volunteers (close to 10% of the college full-time workforce) and in executing its strategies along the way. The Call Center receives an A- from the formative evaluation of its process. Several issues came to the attention of the study. Erin Lee, who ran the Call Center on a

daily basis, documented the key issues to be quickly aging data and less effective design of the form and questions, among other things (page 19). These two issues must be addressed in future Call Center activities.

In conclusion, there is enough evidence to suggest that call centers may have a place in enrollment management and they are highly worthwhile to be conducted periodically in the future. In order to make the future Call Center more successful, the study makes five (5) recommendations. Chief among them,

It is recommended that a method be found to check the registration status nightly of the list of applicants to be called.

It is recommended that Planning & Research Office (PRO) be involved on the front end to help with a true control and experimental design in order to better gauge the impact and effectiveness of the Call Center.

It is recommended that the Call Center become part of the overall enrollment management strategy of the college. In particular, the Call Center may be further enhanced through more concerted efforts by being part of the enrollment services project.

Initial Evaluation of Cabrillo College Spring 2005 Call Center

(Draft)

Jing Luan
Cathy Summa
Marcy Wieland

BACKGROUND

A call center is a medium with which organizations communicate with their members and with which service providers keep in touch with their customers, mostly through the use of telephones staffed by trained personnel. In higher education, call centers have been used to perform similar functions, with actual tasks varying from fund raising, to enrollment follow-up, to survey research, such as those at Piedmont and Sinclair Community Colleges. Sometimes call centers are named "phone banks”. If done well, call centers may provide a positive impact on enrollment.

Personal touch in an era of a fast paced, impersonal lifestyle can leave a person reached by the call center with a lasting impression. Not the least of which is the power of persuasion inherently present in a call from a college representative compared to receiving a piece of mail. To reinforce this concept, a March 2005 issue of Campus Technologies, an article entitled “Getting Personal” discussed strategies being implemented at various institutions to boost their enrollment. Ferris State, for example, attributed the increase of 2,327 students up from 9,495 a couple of years ago to customized recruiting. After providing many examples, the article stated that “Campuses…to interact with potential students have reported success in meeting their enrollment goals”. In the University Business magazine published in the same month, the article “A New Definition of Marketing” discussed the concept of engaging students through organizational as well as departmental marketing efforts.

Spring 2005 marked another round of community college fee increases enacted by the California legislature. Past fee increases had proven to negatively impact enrollment. For example, one study conducted by the college’s Planning & Research Office (PRO) showed that for every fee increase, there had been a corresponding drop in enrollment. A $13 increase would effect a 6% drop in headcounts, which translates into hundreds of students not enrolling. In confirmation of this, in 2004 the Chancellor’s Office for the 109 California’s community colleges estimated that system-wide annually a total of 142,500 students would be “lost” due to such an increase.

In anticipation of a potential enrollment dip, Cabrillo College’s Marketing & Communications department, with assistance from several areas of the college, helped form the enrollment fee task force with the goal to maintain, if not increase, the spring 2005 enrollment. Among many areas identified as worthy of improvement, the issue of low registration rate from a large pool of applicants rose to the top. Although a majority of the applicants would register, thousands may never proceed beyond turning in their applications. Back in spring 2004, a total of 1,836 applicants out of 7,137 did not register for classes. Therefore, one of the strategies identified by the taskforce was to direct Call Center outreach activities to those who have applied, but have not registered in spring 2005. As of January 18th 2005 roughly a month before the spring census day, a total of 2,649 applicants were identified. Between January 10th to January 30th, 2005, Cabrillo College with the generous donation of time from 50 volunteers made calls to a list of applicants at the beginning of the spring 2005 semester.

The purpose of the Call Center was to maintain or increase enrollment for spring 2005. The specific objectives of the Call Center effort were the following:

- Primary: to remind and encourage students to register
- Secondary: to help students resolve problems stopping them from registering.
- Tertiary: to gather data about registration problems and identify any trends.

FOUR KEY QUESTIONS ADDRESSED

This study addresses four (4) key questions in summative and formative evaluation of the spring 2005 Call Center effectiveness. The general term of effectiveness used here means to include Yield Rate, Productivity, and Process Improvement. Yield refers to the number of applicants who have become registrants as a result of the Call Center's efforts. Productivity refers to the average units taken by these registrants. Process Improvement is a type of formative evaluation for the purpose of identifying good practices and weakest links in the entire Call Center effort. An additional question for designing a predictive model by pre-classifying future applicants into groups scaled according to their likelihood to register is also explored by the study. A predictive model would reduce the cost of Call Center by identifying those who are less or least likely to register, so that calls are more focused and targeted.

Specifically, this study addresses four questions.

Question One (Yield): How many of the applied-but-not-registered applicants became registrants as a result of being reached by the Call Center.

Question Two (Productivity): What are the average units taken by the registrants as compared to other registrants who were not directly reached by the Call Center?

Question Three (Predictive Modeling): How many applicants can be predicted to be less likely to register so that the Call Center can concentrate on these applicants?

Question Four (Process Improvement): Are there areas and practices in the Call Center process that may be improved?

DATA SOURCES

Computing Resources (CR) provided lists of applicants who applied but had not registered for select dates based on request from the Call Center volunteers. CR also provided summary counts of applicants for both current semester and historical spring 2004 semester. Call Center provided feedback data in the form of notes taken by the Call Center volunteers. Planning and Research Office (PRO) conducted data matching where possible prior to conducting statistical analyses.

DESIGN AND METHODS

In order to answer all four (4) questions, this study employed a variety of methods and tools. The study adopted a post facto control and experimental design for seeking answers to Question One. Chi-square statistics, regression equations, neural networks, and Classification and Regression Tree (C&RT) have been used by the study for appropriate situations. Data warehousing and SQL (structured query language) technologies directly supported the datasets merging and querying tasks.

The list of 2,949 students was produced by programmers of the Computer Resources department through querying Datatel, an ERP system at the college. The results of the Call Center were notes and check marks provided by the volunteers on survey forms during their calls. These data were hand coded into an Excel worksheet that was imported into Brio Query, a Business Intelligence (BI) tool for the purpose of querying and pivoting variables (building various reports). Most of the answers to Question One and Question Two are provided by Brio Query, assisted by Excel and SPSS (another statistical analysis tool). For Question Four of predicative modeling, the study utilized data mining and a tool called Clementine, a leading industrial strength Business Analytics (BA) application.

FINDINGS

The answer to Question One was obtained through two separate steps. The first step examined the differences in registration rates between a control group and an experimental group. The experimental group would be the group that had the presence of a call center and the control group had not. To compute specific yield rates, which was the second step, required those in the experimental group who were identified to be those who applied but had not yet registered for classes for the Call Center to contact. Not all applicants could be contacted by the Call Center.

Step One - Overall Effect Of The Presence Of Call Center

As the first step in addressing Question One, the study made refinement to the original pair of a control group and an experimental group by splitting them further. The rationale is as follows. As mentioned earlier, all those who turned in their applications as of spring 2004 semester census date (February 23, 2004) became a pseudo control group because no Call Center activities took place in that semester. All those who had their applications on file as of spring 2005 semester census date (February 22, 2005) were the experimental group. The Call Center only functioned for a brief period of time, a month before the start of the spring 2005 semester and the college continued to receive applicants since the lists of applicants were extracted for the calls. This has provided a good opportunity to examine the registration rates with and without the Call Center in the same semester. Therefore, applicants in both groups were then split by a specific date. For spring 2005, the date of January 19th was chosen because none of the applicants who turned in their applications after the 19th of January were contacted. This group is called "Pool A". January 19th 2005 was 31 days before the census day of spring 2005. For spring 2004, the date of January 20th 2004 was chosen (31 days before spring 2004 census date). This group is called "Pool B". Hypothetically speaking, if Call Center had no effect, then the rates for Pools A and Pools B in their respective semesters should be very similar.

The following table presents the rates of registration for the control (Pools A & B) and experimental groups (Pools A & B).

Table 1: Registration Rates by Treatment & Control Groups

		Spring 05		Spring 04
Pool A	Call Center Available	**71%**	No Call Center	**74%**
Pool B	No Call Center	**66%**	No Call Center	**74%**
		Diff = 5%		Diff = 0%

Note:

1) Pools A contain all applicants who turned in their applications as of spring census date. For Pool A from spring 2005, a sub-group of the applicants, namely those who applied but had not registered as of the date of January 19, 2005, was the group contacted by the Call Center and consequently discussed in detail in the study.
2) Online application was made available for the first time in spring 2005. A great number of students utilized it to apply online. This is a key difference other than a Call Center available to both Pool A and Pool B in spring 2005.

Table 1 shows that those spring 2005 students in Pool A had a higher registration rate than Pool B. The difference is 5%. Both the equivalent pools of students in spring 2004 showed no change in their registration rate.

Is the observed 5% difference statistically significant? The study turned to Chi-square analysis for answers. The following output from Chi-square analysis showed a high level of significance.

Figure 1: Chi-square output

$\chi^2 = \sum (O\text{-}E)^2/E =$ 24.52
degrees of freedom= 1
p= 0.000001

The Chi-square analysis indicates that the observed 5% difference was statistically a significant event for the registration rate for those who applied but had not registered in spring 2005. The occurrence of a difference of 5% purely by chance is deemed to be one in 100,000th, or very unlikely.

It should be noted that the control and experimental groups were established after the fact. This is often practiced in education research due to the ethical concerns about intentionally denying access to services in a true control/experimental design. Further, several changes of circumstances existed for both the control and experimental groups. For example, more time was given to the applicants in spring 2004 to collect their applications than those in spring 2005 due to the onset date of research, which was March 8th, 2005. About 2% fewer sections were offered in spring 2005, which may have caused some students to have not been able to find the classes of their choice. Readers of this study need to have these caveats in mind when interpreting the 5% difference in registration rate.

Although the 5% difference is considered statistically significant, in order to completely answer the question on the yield rate the next step is to look at the actual number of yields.

Step Two - Computing Specific Yield Rates

The applicants under study have been categorized into several distinct groups. Those who spoke with the Call Center staff directly, and said they were going to register, were in the Promised group. Those who received a voice mail message from the Call Center volunteers were in the Left Msg group. Those whose phones never answered were in the Not Accessible group. Those who spoke Spanish only were in the Spanish Spk group. Those who provided a wrong number on their applications were in the Wrong No. group. Those whose phone numbers had an area code other than 831 were in the Out of Area group.

Of the applicants who had not registered by January 19, 2005, 370 of them spoke with the Call Center staff directly and said they would register. Eventually, 194 of them were found to have registered as of spring 2005 semester census day, thus producing a yield rate of 52% for the Promised group. Across all categories of applicants in Table 2, this is by far the highest yield rate. The next group that had the highest yield rate (48%) was those who received a voice mail message from the callers.

The study removed all cases from all categories if they were found to have a registration date prior to Jan 19th 2005. This will help with making sure that the subjects under study have not been included in error. Secondly, the study went through the actual survey forms filled out by the Call Center volunteers and paid particular attention to those in the Promised. The purpose of examining the actual feedback from the applicants was to get a sense of the reasons of those 194 direct yields. Many of them stated reasons such as "not clear on what to do next", "have not gotten the time", "procrastinating". Many were thankful that they got the call. It was clear that they indeed may not have registered if they did not receive the calls.

Table 2: Yield Rates by Applicant Types

	Applicants*	Yield	Rate**
Promised	370	194	52%
Left Msg.	842	400	48%
Not Accessible	259	120	46%
Spanish Spk	12	5	42%
Wrong No.	288	115	40%
Out of Area	86	26	30%

*Applicants column has been revised by removing those who had already registered by January 19, 2005.

** Rate refers to the number of actual registrants (yields) from within each category of the applicants. This way the rate can be clearly computed to indicate the yield.

Aside from those who received a voice mail message from the callers, those who registered without speaking directly with the Call Center volunteers (or without being reached by the callers) can be regarded as those who enrolled in classes of their own volition; therefore, subjects with no contact from Call Center (without receiving any treatment). They were the Not Accessible, Wrong No. and Out of Area groups. After collapsing the above six (6) application types (Table 3) into only three, the yield rate of the applicants for the Promised group is ranked 11% higher than the All Other group and is still 3% higher than all others after it is combined with the Left Msg group (Promised & Msg).

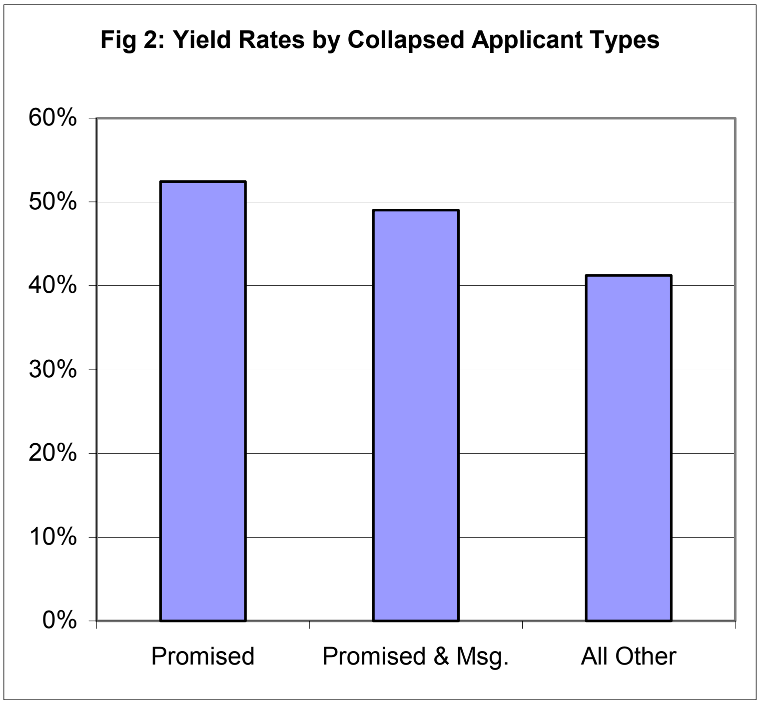


Table 3: Yield Rates for Different Types of Applicants' Categories

	Applicants*	Yield	Rate
Promised	370	194	52%
Promised & Msg.	1212	594	49%
All Other	633	261	41%

*Applicants column has been revised by removing those who had already registered by Jan 19, 2005.

Legend:

Promised denotes those who spoke to the callers directly and promised to register. Promised & Msg denotes those of the Promised group as well as those who received a voice mail message. All Other denotes those who were not reachable, wrong number, and out of the area (who are likely distant ed or non-residents).

PRODUCTIVITY

Have those 194 students who spoke with the Call Center callers directly and eventually enrolled (the direct yield group) actually helped with the college's FTES? Recall this is Question Two on Productivity. This question is answered by calculating the average units attempted by those 194 students. The following table shows the average units attempted by those 194 direct yields and others.

Table 4: Average Units Taken by the Direct Yields and All Others

	AvgUnits
Promised	6.1
Promised & Msg.	5.8
All Other	6.3

Legend:

Promised denotes those who spoke to the callers directly and promised to register. Promised & Msg denotes those of the Promised group as well as those who received a voice mail message. All Other denotes those who were not reachable, wrong number, and out of the area (who are likely distant ed or non-residents).

The difference shows that those who spoke directly with the Call Center staff on average took fewer units than the All Other group by a fraction of a unit. The productivity of the direct yields, as measured by the average units, is similar with those who registered of their own volition.

Nonetheless, the 194 direct yields have collectively enrolled in over 1,000 units.

PREDICTIVE MODELING

Question Three: How would predictive modeling help with identifying among the future applicants, who is less likely to register to help better focus calls made by the Call Center?

To answer this question, the study first used a Venn diagram to group the 2,949 applicants based on their presence in the college's MIS data warehouse and the spring 2005 census database. A Venn diagram, developed by John Venn, a logician, is an effective way to visually identify subgroups (also called sets in Algebra) of any populations, particularly when there is a great amount of overlapping. Venn diagrams have found use in many fields, such as database design, combinatorics (Ruskey, 1997), and assessment of learning (Dixon, 2005).

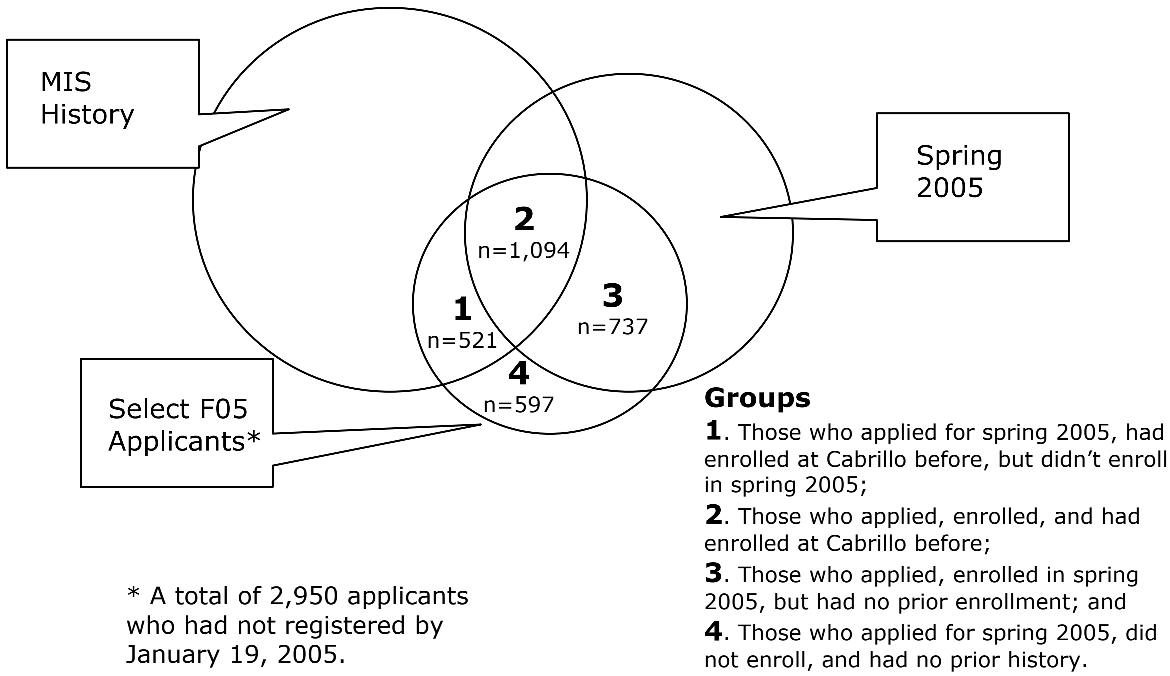
The following Venn diagram (Figure 3) indicates the overlapping of the population of 2,949 applicants with the colleges' MIS historical data warehouse and the spring 2005 census database. Four (4) distinct groups are therefore clearly visible. They are:

Group 1: Those who applied for spring 2005, had enrolled at Cabrillo before, but didn't enroll in spring 2005 (n=521);
Group 2: Those who applied, enrolled, and had enrolled at Cabrillo before (n=1,094);
Group 3: Those who applied, enrolled in spring 2005, but had no prior enrollment (n=737); and

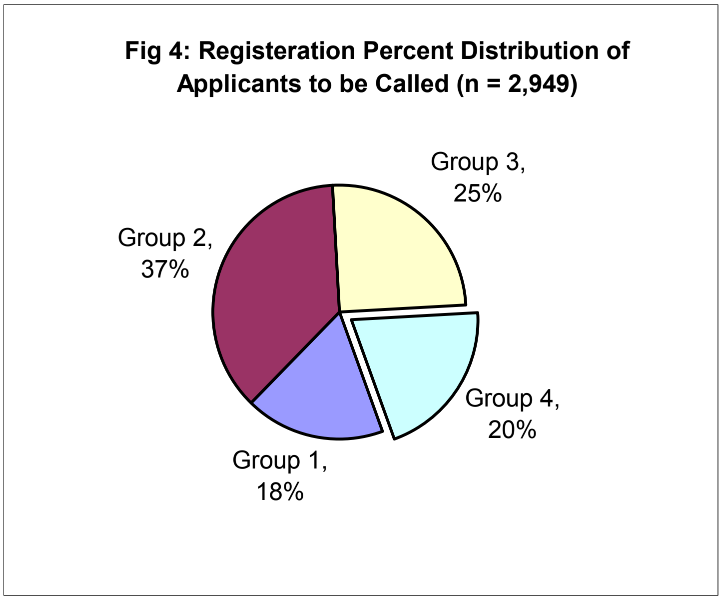
Group 4: Those who applied for spring 2005, did not enroll, and had no prior history (n=597).

These groups helped the rest of the analysis by making it possible to focus on each of them while drilling down to its population details. They guided the rest of the study and are frequently referenced.

Figure 3: Venn Diagram of Four Groups of the 2,949 Applicants



Since Venn diagrams do not display data in proportion to their distributions in a dataset, a pie chart below will correct that by making the distributions adjusted to their appropriate scale. First, for those who eventually registered, the majority (62%, Groups 2 and 3) of the 2,949 applicants eventually registered as of spring 2005 census time. At least a quarter (25%) of the applicants was new to Cabrillo because no prior academic records existed for them in the college MIS historical data warehouse going back 15 years.



For those who never registered, regardless of being reached by the Call Center or not, over a third (38%, Groups 1 and 4) of the 2,949 applicants on the call list did not eventually register, but half of them (Group 1) had attended Cabrillo before. The other half, or 20% of the 2,949 applicants, had never been to Cabrillo.

Overall, 45% (Groups 3 and 4, n=1,334) of the 2,949 applicants had never been to Cabrillo before. It is quite unique to have so many of the potential "new" students among the 2,949 applicants to be called by the Call Center. Was there a reason for a disproportionate number of applicants who had never been to Cabrillo to be slow in registering for classes? Spring 2005 enrollment statistics showed that a total of 2,741 students enrolled were new students. Hypothetically, the number of new students could have been 3,338 (2,741 + 597 of Group 4). In other words, 17.8% (597/3,338) were missing from the new students pool.

The pie chart (Figure 4) seems to indicate that the 2,949 applicants almost had an equal ¼ chance in falling into any of the four groups. Overall, having been to Cabrillo seemed to increase the chance of registering for classes (37%, Group 2). For those who had never been to Cabrillo, their chances of registration were about 50/50. This means that the outcomes of the applicants are really a set of four: those who had been to Cabrillo but did not register and those who did register; those who were new to Cabrillo and registered; and those who did not register and their prior background information is unknown.

The following 5 charts (figures) and tables display background characteristics of the groups identified in the Venn diagram. However, Group 4 is not in any of the analysis due to lack of data.

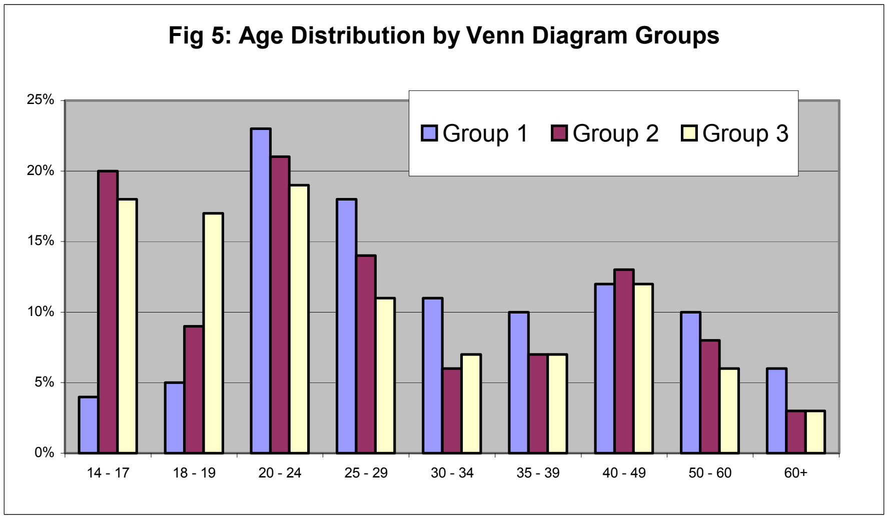


Table 6: Age Distribution by Venn Diagram Groups

	Group 1		Group 2		Group 3	
	Cnt	%	Cnt	%	Cnt	%
14 - 17	23	4%	216	20%	128	18%
18 - 19	24	5%	93	9%	127	17%
20 - 24	119	23%	231	21%	136	19%
25 - 29	96	18%	149	14%	79	11%
30 - 34	58	11%	70	6%	50	7%
35 - 39	54	10%	72	7%	52	7%
40 - 49	64	12%	137	13%	87	12%
50 - 60	53	10%	87	8%	45	6%
60+	29	6%	37	3%	25	3%
Total/Avg	520	100%	1092	100%	729	100%

Figure 5 and Table 6 above show the distribution of age ranges across the three Venn diagram groups.

Overall, the age of students in Group 1 (those applicants who had been at Cabrillo before but did not eventually register) was older than the other two groups. Compared to Groups 2 and 3, Group 1 had fewer students younger than 20. The reverse is true for Groups 2 and 3.

Students ages 17 and below or 19 and below as shown in Figure 5 are likely concurrently enrolled students. For Group 2 (the group of applicants who had taken classes at Cabrillo and registered), there were fewer students in the age range of a recent high school graduate (18-19) compared to Group 3 (new applicants who had never been to Cabrillo). Comparing Group 2 to Group 3, fewer students in Group 2 were from the age range of 18 to 19. The missing ones may have been recent high school graduates who had decided to move on following their study at Cabrillo.

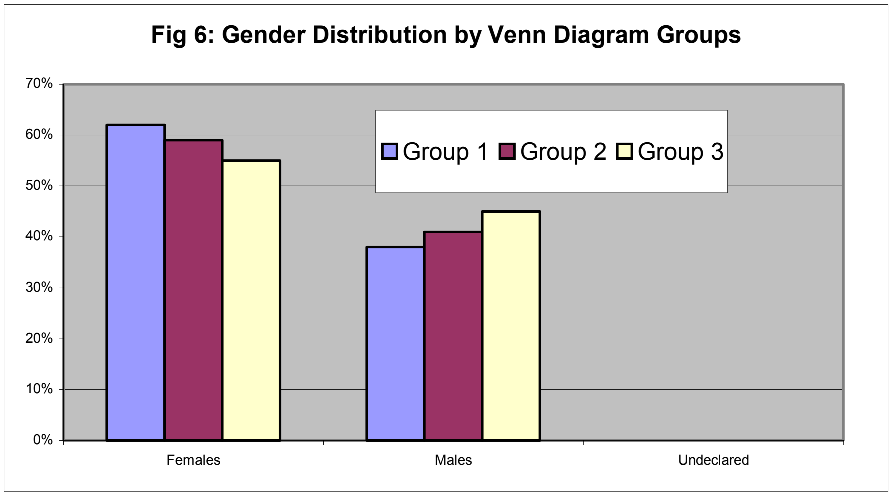


Table 7: Gender Distribution by Venn Diagram Groups

	Group 1		Group 2		Group 3	
	Cnt	%	Cnt	%	Cnt	%
Females	322	62%	645	59%	407	55%
Males	197	38%	446	41%	329	45%
Undeclared	2	0%	3	0%	1	0%
Total/Avg	520	100%	1092	100%	729	100%

Figure 6 and Table 7 above show the distribution of gender across the three Venn diagram groups.

Across the three groups, gender seemed to have an opposing trend. More females were in Group 1, less in Group 2 and much less in Group 3, but the reverse is true for males. New applicants (Group 3) tended to be male. Applicants who had been to Cabrillo and had not registered tended to be female.

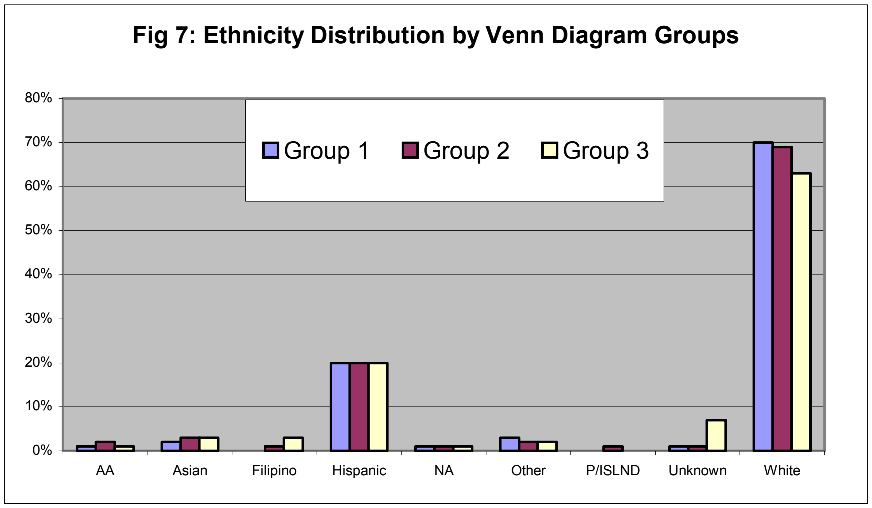


Table 8: Ethnicity Distribution by Venn Diagram Groups:

	Group 1		Group 2		Group 3	
	Cnt	%	Cnt	%	Cnt	%
AA	7	1%	19	2%	9	1%
Asian	13	2%	28	3%	22	3%
Filipino	2	0%	11	1%	19	3%
Hispanic	102	20%	224	20%	147	20%
NA	7	1%	15	1%	8	1%
Other	17	3%	19	2%	14	2%
P/ISLND			7	1%	1	0%
Unknown	7	1%	13	1%	52	7%
White	366	70%	758	69%	465	63%
Total/Avg	521	100%	1094	100%	737	100%

Figure 7 and Table 8 above show the distribution of ethnicities across the three Venn diagram groups.

There is no major difference across major ethnic minorities among all three groups of applicants. There appeared to have fewer White students in Group 3 (new applicants without Cabrillo experience) while there is an increase in the Unknown category in Group 3. Research has shown that most of the students in “unknown” or “unreported” categories tend to be White students.

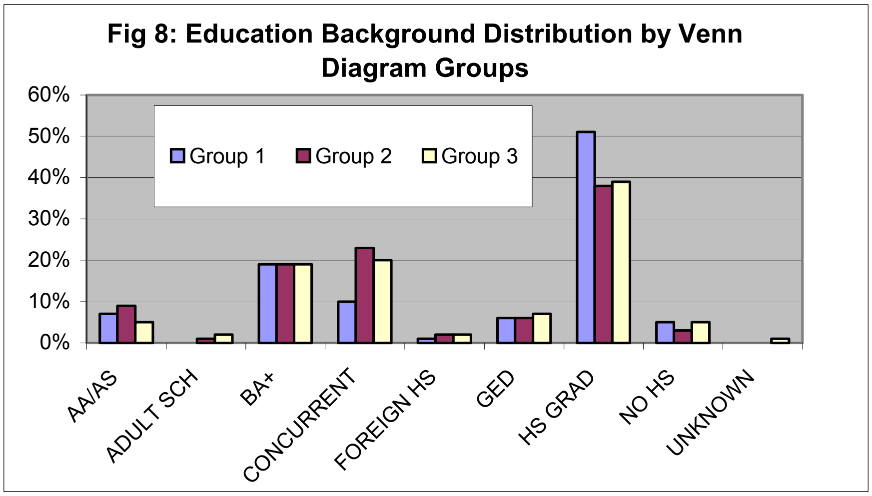


Table 9: Education Background Distribution by Venn Diagram Groups

	Group 1		Group 2		Group 3	
	Cnt	%	Cnt	%	Cnt	%
AA/AS	37	7%	96	9%	35	5%
ADULT SCH			6	1%	13	2%
BA+	98	19%	204	19%	137	19%
CONCURRENT	52	10%	252	23%	149	20%
FOREIGN HS	6	1%	22	2%	13	2%
GED	32	6%	62	6%	54	7%
HS GRAD	268	51%	415	38%	289	39%
NO HS	26	5%	34	3%	36	5%
UNKNOWN	2	0%	3	0%	11	1%
Total/Avg	521	100%	1094	100%	737	100%

Figure 8 and Table 9 above show the distribution of education background across the three Venn diagram groups.

There were fewer concurrently enrolled students in Group 1 (those applicants who had been to Cabrillo, but never registered) compared to other groups. This is a validation of the observations made about their age. More of them in Group 1, on the other hand, were high school graduates, but not necessarily recent high school graduates.

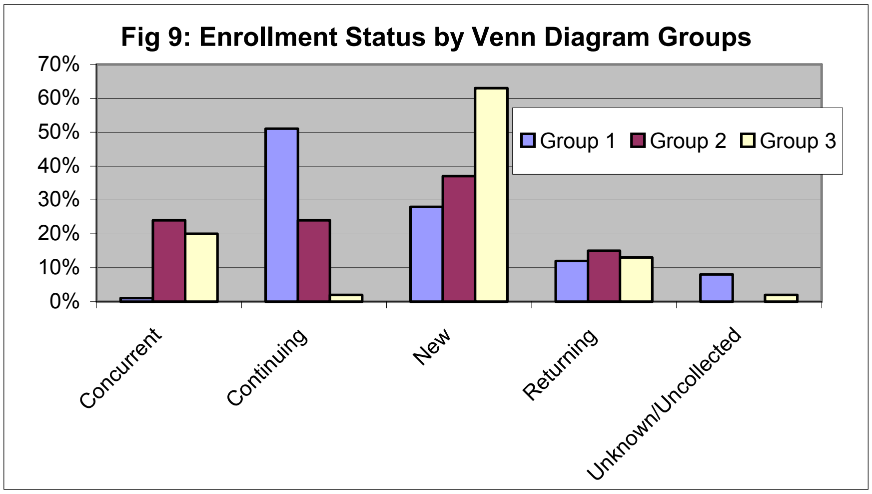


Table 10: Enrollment Status by Venn Diagram Groups

	Group 1		Group 2		Group 3	
	Cnt	%	Cnt	%	Cnt	%
Concurrent	5	1%	258	24%	151	20%
Continuing	264	51%	259	24%	13	2%
New	145	28%	409	37%	462	63%
Returning	64	12%	164	15%	97	13%
Unknown/Uncollect	43	8%	4	0%	14	2%
Total/Avg	521	100%	1094	100%	737	100%

Figure 9 and Table 10 above show the distribution of enrollment status across the three Venn diagram groups.

The largest portion of Group 1 was those who were continuing students when they applied. The largest portion of applicants in Group 2 was those who were new. The largest portion of applicants for Group 3 was those who were new as well.

Distributions of enrollment status for the three groups of applicants were generally very diverse. Very few in Group 1 were concurrently enrolled students. Very few in Group 2 were continuing students when they applied.

The above analysis by five (5) select background variables of demographics and academic status help develop an impression of the different characteristics of the applicants in Groups 1, 2 and 3. However, the impression is at best a fuzzy one, not accurate or evidential to help classify individual future applicants into respective groups. The exercise of examining these background variables using the method above can go on forever and there can be infinite number of tables and charts should one decide to cross tabulate the groups by two (2) or more variables.

The rest of the study, in response to Question Three, employed data mining technique. Specifically, the study used Neural Net and Classification & Regression Tree (C&RT) nodes. Both are predictive modeling nodes based on artificial intelligence and machine learning. The predicative modeling nodes, to state simply, has the ability to study known cases in order to make predictions onto unknown cases. The unknown cases would be incoming applicants for a future semester. Along the way, data mining also provides a number of sophisticated ways of examining and describing the data. For detailed rationale and design of using data mining for this study, please consult Appendix A).

Not all cases have been used for the predictive modeling. First, hundreds of applicants spoke directly with the Call Center staff, which could have contaminated the subjects because of the extra intervention they received. The direct interaction with the Call Center staff would artificially make the likelihood of registering for classes higher, particularly such an interaction has been proven to be effective in generating enrollments. Therefore, the study removed these 370 applicants. Secondly, the study removed the entire Group 4, because there was no background data for this group of applicants who never registered. Thirdly, Group 3 (those who had no prior records at Cabrillo but eventually registered) presented another dilemma. The most significant marker for this group is that they had no prior records, which would trick the algorithm to put too much weight on this fact that could overwhelm all other variables. Therefore, the entire group was removed as well. The remaining 1,413 applicants came from Group 1 (n=463) and Group 2 (n=950). The outcome variable, S05Regd, is binary: registered “S05” and not registered “NoS05”.

Between applicants registering for classes and not registering for classes, of interest to the study is those who did not register. The reason is simple. If data mining algorithms can predictively identify those who are less likely to register, then the Call Center staff can concentrate on these applicants, which is a far effective use of time and resources. Defined as Between Group Accuracy, all the applicants who did not register (Actual) should be “predicated” as not having registered (Predicted). The idea is for the data mining algorithms to first examine all the variables associated with each applicant’s registration outcome to learn the rules and then apply the rules to similar cases to predict their registration outcome. Since the focus of the predictive modeling is on those who would not register, the level of Between Group Accuracy should be higher than 85%, which means out of all the applicants predicted to be less likely to register, it would get 9 out 10 of them correct.

The Neural Net node was run first with the following Sensitivity Statistics:

Table 11: Relative Importance of Inputs of Neural Net Node

Variable Name	System Variable Names	Sensitivity
Enrollment Status	Group of EnrlStatAll	0.40344
High School Origin	HighSchAll	0.33039
Age (in ranges)	AgeRange2005	0.307335
Ethnicity	Group of EthnicityAll	0.30281
Education Background	Group of EdStatusAll	0.249625
Number of Terms Previously Enrolled	CntDTL	0.176079
Gender	GenderAll	0.115245

Note: Neural Net node: 71.2% accuracy; 59 neurons, 1 hidden layer; Quick, seed true.

The sensitivity statistics from the Neural Net node helps the analysts understand the level of importance of input variables that determine an outcome. In this case, the outcome is registered or not registered. The node indicates (Table 11) that the enrollment status is the most important variable and gender is the least important variable.

Yet, the predication accuracy by the Neural Net node is not impressive. The between group predictive accuracy of Neural Net node for those who did not register is low (55.9%), evidenced by the following matrix (Table 12). This represents a slightly better chance than a coin toss. Neural Net node was not used after this discovery.

Table 12: Between Group Accuracy Matrix by Neural Net Node

NN:		Predicted	
Actual		Not Reg'd	Registered
Not Reg'd	Count	259	204
	Row %	**55.9**	44.1
Registered	Count	160	790
	Row %	16.8	**83.2**

C&RT node was used next without any adjustments inside the node to accommodate for misclassification. Misclassification is the equivalent of false positives. It is often used when the analysts choose to err on the side of predicting more cases favoring one outcome over the other, such as not registering for classes, so as to increase the chances of getting more cases of one side (type) correctly predicted at the expense of the other.

In Table 13, the between group predictive accuracy for those who did not register is much better (85.7%) than those who eventually registered (70.4%).

Table 13: Between Group Accuracy Matrix by C&RT Node

C&RT		Predicted	
Actual		Not Reg'd	Registered
Not Reg'd	Count	393	70
	Row %	**84.9**	15.1
Registered	Count	344	606
	Row %	36.2	**63.8**

The rest of the analysis utilized a decision tree graph and rule sets from the C&RT node.

Figure 10: Decision Tree (Partial) from C&RT

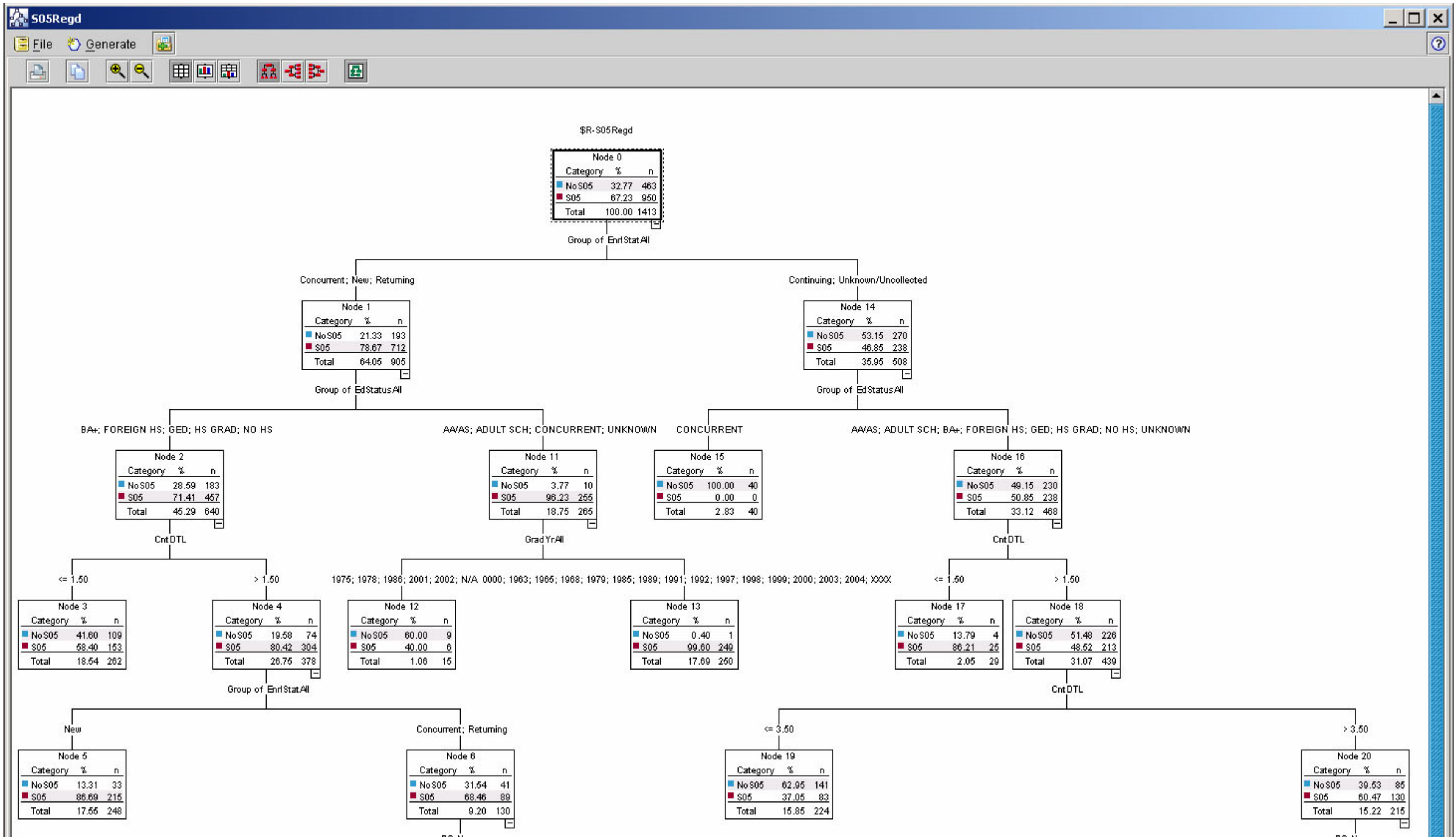


Figure 10 is a partial screenshot of a binary decision tree produced by the Classification and Regression Tree (C&RT) algorithm. The tree is six (6) branches deep from the first node of registration status ($R-S05Regd), but only the first four (4) top branches can fit on a landscape page in this report. The top branches typically contain variables more influential to the outcome than the rest toward the bottom.

In the decision tree graph (Figure 10), applicants' latest Enrollment Status seemed to be most important, therefore, it became the first split. Those who were Concurrently Enrolled, New or Returning applicants (n=905) were split from those who were Continuing or with unknown status (n=508). Of those 905 applicants, 21.3% (n=193) did not register and 78.7% of them did. For the sake of understanding the tree graph, let's focus on these 193 applicants who did not register. In the next split that occurred on the variable of Education Background, 183 were left. They were those who had a BA, or a Foreign High School Diploma, or a GED, and those who either graduated or did not graduate from high school. The next split took place on the variable of number of terms enrolled at Cabrillo prior to applying for college (CntDTL), 109 applicants who had less than 1.5 total terms enrolled were left. They reside in the terminal node, which means no further split occurred. Therefore, tracing the above splits, it can be reasonable stated that the 109 applicants who did not register were those who had a very short or no prior attendance at the college, who had a BA, or sort of high school diploma, regardless of their high school graduation status, and were either concurrently enrolled, or a new or returning to college after a brief stop out.

To identify these patterns manually can be cumbersome. Fortunately the data mining algorithm has already combed through the tree diagram and produced a list of the patterns, called rules. There are a total of 6 rules for those who did not register and another set of 6 rules for those who did. If more variables were introduced to the algorithm, the number of rules would likely increase. The first rule reads as follows:

```
Rules for No Registration - contains 6 rule(s)
    Rule 1 for No Registration (262, 0.584)
        if Group of Enrollment Status in [ "Concurrent" "New" "Returning" ]
        and Group of Education Background in
            [ "BA+" "FOREIGN HS" "GED" "HS GRAD" "NO HS" ]
        and Number of Terms Ever Enrolled <= 1.500
        then No Registration

    Rule 2 for  No Registration (36, 0.528)
        if Group of Enrollment Status in [ "Concurrent" "New" "Returning" ]
        and Group of Education Background in
            [ "BA+" "FOREIGN HS" "GED" "HS GRAD" "NO HS" ]
        and Number of Terms Ever Enrolled CntDTL > 1.500
        and Group of Enrollment Status in [ "Concurrent" "Returning" ]
        and City Locations in
            [ "Aptos" "Campbell" "Capitola" "Moss Landing" "Mount Hermon" "Santa Cruz"
        "Soquel" "Watsonville" ]
        and Age Range in [ "25 - 29" "35 - 39" ]
        then No Registration
```

The first rule corresponds to the observations made early for the 109 applicants. Because 153 of those who actually registered were also classified into the terminal node with the 109 applicants, therefore, the rule states at the beginning that there were a total of 262 cases and the confidence is only 58%. It is not unusual that the accuracy for individual rules may not be always high. The 4th rule had 40 cases in it and it had a confidence level of 100%. The rules are listed without a hierarchical order. The first rule is as important as the last rule. (See appendix for all 12 rules.) To save paper, let's take a look at the first rule for those who would register.

```
Rules for Registration- contains 6 rule(s)
        Rule 1 for  Registration (248, 0.867)
                if Group of Enrollment Status in [ "Concurrent" "New" "Returning" ]
                and Group of Education Background in
                        [ "BA+" "FOREIGN HS" "GED" "HS GRAD" "NO HS" ]
                and Number of Terms Ever Enrolled > 1.500
                and Group of Enrollment Status in [ "New" ]
                then Registered.
```

There were a total of 248 cases belonging to this rule and its confidence level was 87%.

It is clear through just viewing a few of the rules that there are many different profiles that could determine the final outcome of registering for classes. This proves that there is no one equation that can adequately illustrate the highly various nature behind the registration status of all the applicants. Case in point, only about 40 students fit the 4th Rule for those who did not register. Another 248 of the applicants fit the 1st Rule for those who registered.

The 2,949 applicants, a subset of all the applicants for spring 2005, were those who for some reason had not registered just days before the semester was to start, the analysis of their behaviors and background information should bear in mind that real reasons and the motivation factor were unknown. Some of them may be straddlers or stragglers, others may have a legitimate reason to be slow. As the notes taken by the Call Center staff showed, some applicants were confused by the registration process, some were waiting for their appointment dates, some took a backseat when they realized they had to go through assessment. Quite a few applicants reached by the Call Center said they were moving out of the area, had financial aid issues, or downright unmotivated and "lazy" quoting their own words.

At the time of this study, academic history data is only available for Groups 1 and 2. When background data became available for Group 4, the predictive modeling may be further enhanced.

PROCESS IMROVEMENT

Question Four: What parts of the processes can be identified as practice-worthy and others as areas for improvement? (Process Improvement Question)

The Call Center was well orchestrated and materials well organized. The instructions were clear and adequate training was available to all volunteers. The cell center used staff donated time, so the direct cost of operating it was minimal. The indirect cost to the college was also low, considering the fact that many staff and faculty volunteers made calls in the late afternoon, early evenings and weekends. Periodical reminders and email messages with tips and notices were distributed timely.

The volunteerism spirit was ultra high and alive at Cabrillo. A total of 50 volunteers, or 10% of the entire full-time faculty and staff, took part in calling students. They were office assistants, faculty, deans and the college president.

The Marketing and Communications department, the key organizer of the Call Center, helped identifying the following areas for potential improvement.

- Quickly aging data: The call lists were based on a snapshot of potential students who had applied but not yet enrolled. Frequently, by the time callers made contact with a student on their call list, the student had already enrolled. The longer callers waited to make their calls from the time the list was generated, the more likely they were to reach a registered student. Several of the callers had access to student records in Datatel, and checked on students' registration status before calling them to counter this problem.

- Varying data: In meeting the objective of tracking registration problems, the questions in the call script were found to be too open-ended and liable to be interpreted differently by various callers. In addition, callers had a very limited amount of time on the phone with students and didn't always ask all the questions on the call form. The resulting data sometimes resisted easy categorization when trying to find any glaring trends in registration problems. For the next Call Center, a tighter call form with fewer (and more targeted) questions is recommended.

- Language: For the next Call Center, more Spanish-speaking volunteers should be recruited.

- Availability of help: For the next Call Center, there should be a telephone number to leave with the applicants, enabling them to call back and reach someone with knowledge of their specific problem. In a perfect world, there ought to have one telephone staffed for the duration of the Call Center project by someone with access to all the Call Center information, and the ability to solve admissions and registration difficulties then and there.

- Form Design: Move the categories to the front and move the call logs down, so that callers will not only focus on reaching the students.

DISCUSSIONS

In responding to the four (4) research questions, the study uncovered the following findings. In terms of Call Center effectiveness as defined by yields and significance of the yields (Question One), the study found a statistically significant 5% higher registration rate in a four way post-facto control and experimental analysis. But this significance should also take into consideration a few college initiated changes, although no specific impact to any of the four way analyses was immediately perceivable. For example, the college implemented online application in spring 2005, which increased the total applicants by 12% as compared to the spring a year earlier.

The study found a noticeable difference in registration rate large enough to attribute the increase in registration to the Call Center after examining the different categories of applicants coded for callers. Those who spoke directly with the Call Center staff had at least a 10% increase in registering for classes compared to those who had no contact with the Call Center. Yet in answers to Question Three on Productivity, the increase was found at best to have helped maintain the level of FTES, a productivity measure. However, it can be viewed that without the Call Center, there could have been a perceivable drop in both headcounts and FTES in spring 2005. From a summative evaluation perspective, the Call Center receives a B+.

After removing the applicants who interacted with the Call Center directly, data mining analysis identified six (6) rules for applicants who would register for classes and another six (6) for those who would not. Further, the data mining analysis built a predictive model with an 85% accuracy to predict those who would be less likely to register in a future semester.

The Call Center was a success in mobilizing an army of volunteers and in executing its strategies along the way. The Call Center receives an A- from the formative evaluation of its process. Several issues came to the attention of the study. Erin Lee, who ran the Call Center on a daily basis, documented the key issues to be quickly aging data and less effective design of the form and questions, among other things. These two issues must be addressed in future Call Center activities.

At the turn of the millennium, Cabrillo College conducted a round of Call Center activities. The callers were equipped with the ability to register students while on the call. Close to 300 students out of 2,500 number reached by the Call Center registered this way. With the recent Call Center, there is enough evidence to state that call centers have a place in enrollment management and they are highly worthwhile to be conducted periodically in the future.

RECOMMENDATIONS

This study makes the following five (5) recommendations.

It is recommended that the Call Center be conducted in either fall 2005 or spring 2006 semester.

It is recommended that a method be found for callers to check the registration status nightly of the list of applicants to be called.

It is recommended that PRO be involved on the front end to help with a true control and experimental design in order to better gauge the impact and effectiveness of the Call Center.

It is recommended that the form used by the Call Center be revised so that callers can quickly jot down reasons applicants have not registered and use these reasons to identify assistance.

It is recommended that the Call Center become part of the overall enrollment management strategy of the college. In particular, the Call Center may be further enhanced through more concerted efforts by being part of the enrollment services project.

REFERENCES (Partial):

Borden, R. (2003) Projected Impact of Fee Hike on Cabrillo Fall Enrollment 2003 to 2007. Cabrillo College Publication. Mary 2003.

Chang, L. (2005) A Case Study Applying Data Mining Technology in Modeling and Predicting Admissions Yield in Higher Education. AIR Forum, San Diego, May 2005.

Digby K. E. (1986) A Study of the Factors Which Influence Adult Enrollment in a Technical Institute. ERIC (ED273319).

Fetterer, F. and K. Knowles. 2002, updated 2004. Sea Ice Index. Boulder, CO: National Snow and Ice Data Center. Digital media.

Fluentes-Michel, D., Perry, P. (2003). Student Fees. Presentation at the Study Session of the Board of Governors Meeting of California Community Colleges (http://www.cccco.edu/executive/bog/agendas/attachments%5F0303/14%2Dstudent%5Ffees.doc). March 10-11, 2003.

Hossler, D. (1984) Enrollment Management – an integrated approach. The College Board. New York, New York.

Mena, J. (1998) Data Mining FAQs. DM Review. January 1998.

Rounds, J. C., (1984) Attrition and Retention of Community College Students: Problem and Promising Practices. ERIC (ED242377).

Ruskey, F. (1997). A Survey of Venn Diagrams. The Electronic Journal of Combinatorics. March 2001 (online version), DS#5.

Shachmurove, Y. (?) Applying Artificial Neural Networks to Business, Economics and Finance. (online: http://qudata.com/lib/neuronet/)

Sevier, R. (2005) A New Definition of Marketing – The AMA's update can mean enormous opportunities for higher education. University Business. March 2005.

Tansey, F. (2005) Getting Personal – eRecruiting technologies now enable schools to match their strengths to prospective students' individual expectations. Campus Technologies. March 2005.

Zhao, C. M. & Luan, J (2005) Chapter One (draft): Introduction to Data Mining Applications for Enrollment Management. Data Mining Concepts, Myths & Its Applications in Enrollment Management. Proposed issue for New Directions for Institutional Research.

Web Resources:
http://www.sinclair.edu/stservices/outserv/OutReachServiceLinks/index.cfm
http://www.cpcc.edu/ess/enrollment_management/

Venn use:
http://www.valdosta.edu/~jharris/assessment.html

APPENDIX

(A) DATA MINING PREDICTIVE MODELING RATIONALE AND DESIGN

When one (1) variable, such as gender, is examined by some type of an indicator, it is operating in a one dimensional environment, such as Figure 6. When two (2) variables are examined by the same indicator, it can still be laid out in a one dimensional environment using a table or chart, but becomes difficult for the reader to quickly comprehend the information. When there were two (2) or more variables involved, an invisible 3-dimension is formed. Each variable becomes a vector and where they join is a coordinate. The environment in which these variables exist, if plotted in a 3-D graph, is called a Euclidian hyper sphere.

A 3-dimensional environment with multiple variables intermingled in a myriad of ways is impossible for the human eyes to quickly spot trends or monitor changes. It is precisely such a spatial sphere in which hidden patterns exist that can lead to new information on how the applicants become a registrant. Conventional tools may start to show inadequacies in handling infinite number of coordinates in such a spatial sphere. Even the traditional regression analysis, which computes the residual statistics of multiple variable around a mean to determine their contribution to explain the variance of a phenomenon, is not entire adequate.

For example, there may be a dozen variables going into a regression analysis to identify the key ones that would determine whether or not an applicant would register. The regression analysis may find that among these variables gender, ethnicity, age, location, and GPA would be significant. It then provides a specific value (odds ratio) associated with each variable to determine the likelihood of an applicant's registration status. The equation typically functions as a polynomial model: if the value of a variable changes by one unit, such as one (1) year of age increase for an applicant, the likelihood of registration would change by an X amount. Plugging in an applicant whose age is 18, gender is male, and ethnicity is Asian, out comes the likelihood of his registering for classes.

This likelihood is essentially a quadratic equation following a Sigmoid curve (regression line) with its values ranging from 0 to 1. Every new case is fit into this regression equation. The sizes of the confidence intervals drawn for individual cases are much greater or wider than those drawn for the averages predicted by the regression lines (Chang, 2005). For example, not necessarily every 18 year old, male, Asian applicant would be likely to register. Some 18 year old male Asian applicants may register because of variables A, B, and C, and others C, D and F.

Along with logistic regression, linear regression, have been considered inadequate to address most social science problems and has remained as an introduction to modern regression statistical analysis (Fetters, 2002). Not everything in life is linear. As a matter of fact, life is mostly non-linear with a great preference for cases considered outliers. Approaches to nonlinear data do exist, such as Nonlinear Regression (NLR), but it requires both input and output data to be quantitative. Many data mining algorithms do not have such requirement. The Neural Net algorithm, for example, does not make assumptions about the data, nor does it require data to be coded any particular way (Shachmurove, year??).

Jesus Mena in 1998 observed the differences between statistics and data mining by classifying traditional statistics as a top down approach and data mining bottom up. Mena stated that traditional statistics require the researcher to first formulize a theory in the form of a hypothesis and then query the data to prove or disprove the theory. Data mining on the other hand relies on data to self-organize and may not ever rise up to the level of formulating a theory (Zhao & Luan, 2005). Mena also discussed the need for scaling both the analytical prowess and data file to speed and size that is particularly important in a fast moving business environment. For example, according to Mena, a typical database may contain 100 attributes about customers, which would mean, for any customer, there would be 100 x 99 combinations of attributes to consider. A small task such as classifying customers into high, medium or low groups, would mean 970,200 (100 x 99 x 98) possible combinations. Hypothesis testing, coupled with difficulty in scaling to large, ever changing databases, would make approaches with an academic research bent more of a challenge than a tool to solve problems.

Even though not a perfect tool to master the highly complex and intricate relationships that are typical of the cases for predictive modeling, the weaknesses of traditional regression tools of linear and logistic regressions may well be their strengths. They are widely used to identify key variables most contributive to a given outcome. For this purpose, sophisticated parameter statistics have been developed, such as Wald statistic and Hosmer-Lemeshow goodness of fit. Coupled with carefully monitored assumptions about the data, traditional regression tests are highly useful for conclusions around and about summary statistics, means or variances vs. around and about individual cases as C&RT, Neural Net are. Further, for clinical situations of testing group differences, such as control/experimental, pre-post testing, the entire slate of traditional statistical tests are indispensable.

On the other hand, the strengths of newly developed data mining algorithms, though essential to automation, adaptable to data with any depths of complication, are also the weakest in satisfying the needs filled by traditional statistics. It’s naturally then to conclude that data mining may be well served by applying one or two traditional statistical tests along side neural networks, Bayesian or genetic algorithms, just to confirm and to explain at summary level what is at work.

Therefore, to conduct unitary record level predictive modeling is to use data mining tools to help uncover the hidden patterns among all possible enumerations of the variables in a 3-D spatial environment. Data mining, a new frontier in data analysis and known for its ability to use multiple analytical algorithms to process data in a Euclidian spatial sphere therefore is the tool of choice for answering Question Three. One key capability of data mining is to use artificial intelligence, Bayesian, and machine learning techniques to study known cases in order to develop the ability to make predictions onto unknown cases. The unknown cases would be incoming applicants for a future semester. Along the way, data mining also provides a number of sophisticated ways of examining and describing the data. Because multiple algorithms are used, multiple ways are also available for analyzing the data. Some may use decision trees (a graphical representation of the hierarchy of variables to show their lineage of contribution to an outcome, such as registration), others use induction based rule sets that would give a separate set of rules for those who register and another set for those who would not. Both the decision trees and the rule sets are far more powerful in describing the data better than the charts and tables produced on preceding pages.

The study chose Clementine as the data mining tool. The reason for selecting Clementine is based on its ability to directly interface with static or live relational databases, to calculate new fields using GUI (graphical user interface) guided nodes, to convert transactional data files into analytical data files, and to allow infinite number of scenarios to be built and examined using its 16 modeling algorithms, including two (2) text mining nodes. All analyses are conducted inside one data stream, which makes it much easier for cross-validation, interpretation, replication and documentation. The following screenshot illustrates the "data stream" built within Clementine for the entire study, including the nodes used for calculating new fields (variables).

Since Brio Query queried the datasets and produced a data cube containing most of the needed data elements that lent themselves readily as input variables for the data mining tasks, Clementine directly imported a tab delimited text files from Brio Query as its data source.

Not all cases have been used for the predictive modeling. First, hundreds of applicants spoke directly with the Call Center staff, which could have contaminated the subjects because of the extra intervention they received. The direct interaction with the Call Center staff would artificially make the likelihood of registering for classes higher, particularly such an interaction has been proven to be effective in generating enrollments. Therefore, the study removed these 370 applicants. Secondly, the study removed the entire Group 4, because there was no background data for this group of applicants who never registered. Thirdly, Group 3 (those who had no prior records at Cabrillo but eventually registered) presented another dilemma. The most significant marker for this group is that they had no prior records, which would trick the algorithm to put too much weight on this fact that could overwhelm all other variables. Therefore, the entire group was removed as well. The remaining 1,413 applicants came from Group 1 (n=463) and Group 2 (n=950). The outcome variable, S05Regd, is binary: registered "S05" and not registered "NoS05".

Between applicants registering for classes and not registering for classes, of interest to the predictive modeling is those who did not register. The reason is simple. If data mining algorithms can predictively identify those who are less likely to register, then the Call Center staff can concentrate on these applicants, which is a far effective use of time and resources. Compared to knowing who would NOT registers, there is less business advantage to know who are likely to register. Since the predicted outcome is registered and not registered, it is hoped the predictive accuracy for those who did not register should be higher than those who did, if not equal. This is called "between group" accuracy.

(B) COMPLETE RULESETS DEVELOPED BY C&RT

Rules for NoS05 - contains 6 rule(s)

Rule 1 for NoS05 (262, 0.584)
if Group of EnrlStatAll in ["Concurrent" "New" "Returning"]
and Group of EdStatusAll in ["BA+" "FOREIGN HS" "GED" "HS GRAD" "NO HS"]
and CntDTL <= 1.500
then NoS05

Rule 2 for NoS05 (36, 0.528)
if Group of EnrlStatAll in ["Concurrent" "New" "Returning"]
and Group of EdStatusAll in ["BA+" "FOREIGN HS" "GED" "HS GRAD" "NO HS"]
and CntDTL > 1.500
and Group of EnrlStatAll in ["Concurrent" "Returning"]
and PO Name in ["Aptos" "Campbell" "Capitola" "Moss Landing" "Mount Hermon" "Santa Cruz" "Soquel" "Watsonville"]
and AgeRange2005 in ["25 - 29" "35 - 39"]
then NoS05

Rule 3 for NoS05 (15, 0.6)
if Group of EnrlStatAll in ["Concurrent" "New" "Returning"]
and Group of EdStatusAll in ["AA/AS" "ADULT SCH" "CONCURRENT" "UNKNOWN"]
and GradYrAll in ["1975" "1978" "1986" "2001" "2002" "N/A"]
then NoS05

Rule 4 for NoS05 (40, 1.0)
if Group of EnrlStatAll in ["Continuing" "Unknown/Uncollected"]
and Group of EdStatusAll in ["CONCURRENT"]
then NoS05

Rule 5 for NoS05 (224, 0.629)
if Group of EnrlStatAll in ["Continuing" "Unknown/Uncollected"]
and Group of EdStatusAll in ["AA/AS" "ADULT SCH" "BA+" "FOREIGN HS" "GED" "HS GRAD" "NO HS" "UNKNOWN"]
and CntDTL > 1.500
and CntDTL <= 3.500
then NoS05

Rule 6 for NoS05 (160, 0.531)
if Group of EnrlStatAll in ["Continuing" "Unknown/Uncollected"]
and Group of EdStatusAll in ["AA/AS" "ADULT SCH" "BA+" "FOREIGN HS" "GED" "HS GRAD" "NO HS" "UNKNOWN"]
and CntDTL > 1.500
and CntDTL > 3.500
and PO Name in ["" "Ben Lomond" "Davenport" "Freedom" "Los Gatos" "Salinas" "Santa Cruz" "Saratoga" "Soquel" "Watsonville"]
then NoS05

Rules for S05 - contains 6 rule(s)

Rule 1 for S05 (248, 0.867)
if Group of EnrlStatAll in ["Concurrent" "New" "Returning"]
and Group of EdStatusAll in ["BA+" "FOREIGN HS" "GED" "HS GRAD" "NO HS"]
and CntDTL > 1.500
and Group of EnrlStatAll in ["New"]
then S05

Rule 2 for S05 (20, 0.95)
if Group of EnrlStatAll in ["Concurrent" "New" "Returning"]
and Group of EdStatusAll in ["BA+" "FOREIGN HS" "GED" "HS GRAD" "NO HS"]
and CntDTL > 1.500
and Group of EnrlStatAll in ["Concurrent" "Returning"]
and PO Name in ["Aromas" "Ben Lomond" "Boulder Creek" "Castroville" "Felton" "Freedom" "Los Gatos" "Redmond" "Scotts Valley" "Seaside"]
then S05

Rule 3 for S05 (74, 0.716)
if Group of EnrlStatAll in ["Concurrent" "New" "Returning"]
and Group of EdStatusAll in ["BA+" "FOREIGN HS" "GED" "HS GRAD" "NO HS"]
and CntDTL > 1.500
and Group of EnrlStatAll in ["Concurrent" "Returning"]
and PO Name in ["Aptos" "Campbell" "Capitola" "Moss Landing" "Mount Hermon" "Santa Cruz" "Soquel" "Watsonville"]
and AgeRange2005 in ["18 - 19" "20 - 24" "30 - 34" "40 - 49" "50 - 60" "60+"]
then S05

Rule 4 for S05 (250, 0.996)
if Group of EnrlStatAll in ["Concurrent" "New" "Returning"]
and Group of EdStatusAll in ["AA/AS" "ADULT SCH" "CONCURRENT" "UNKNOWN"]
and GradYrAll in ["0000" "1963" "1965" "1968" "1979" "1985" "1989" "1991" "1992" "1997" "1998" "1999" "2000" "2003" "2004" "XXXX"]
then S05

Rule 5 for S05 (29, 0.862)
if Group of EnrlStatAll in ["Continuing" "Unknown/Uncollected"]

and Group of EdStatusAll in ["AA/AS" "ADULT SCH" "BA+" "FOREIGN HS" "GED" "HS GRAD" "NO HS" "UNKNOWN"]
and CntDTL <= 1.500
then S05

Rule 6 for S05 (55, 0.818)
if Group of EnrlStatAll in ["Continuing" "Unknown/Uncollected"]
and Group of EdStatusAll in ["AA/AS" "ADULT SCH" "BA+" "FOREIGN HS" "GED" "HS GRAD" "NO HS" "UNKNOWN"]
and CntDTL > 1.500
and CntDTL > 3.500
and PO Name in ["Aptos" "Boulder Creek" "Capitola" "Castroville" "Scotts Valley"]
then S05

Default: NoS05

(C) CALL CENTER FORM & SCRIPTS

Caller's name: _______________

Caller's extension #: __________

Spring 2005 Student Access Campaign Call Center
Call Form

Student's Name: ____

Student's Datatel #: ____

Student's Telephone #: ____

CALL ATTEMPTS

DATE											
No Answer											
Left msg											
Wrong #											
Call Back											
Completed											

1. Has attempted to register? ☐Yes ☐No
   a. If YES, problems encountered, if any (check all that apply)
      ☐Hawktalk technical or other difficulties (Please explain: ______________________)
      ☐WebAdvisor technical or other difficulties (Please explain: ______________________)
      ☐Delinquent fees
      ☐Residency issue
      ☐Registration holds—one or more of the following: ☐Library ☐Unpaid financial aid ☐Other
      ☐Transcripts not on file for prerequisite
      ☐Course repetition
      ☐Course overlap
2. Is planning to register? ☐Yes ☐No
   a. If NO
      ☐Financial reasons
      ☐Personal circumstances (Please explain: ______________________)
      ☐Planning to attend: ☐other community college ☐CSU ☐UC ☐Private college/other
      ☐Other (Please explain: ______________________)
3. Currently receiving financial aid? ☐Yes ☐No
   a. If NO, planning to apply for financial aid? ☐Yes ☐No

Comments:

(D) LIST OF SPRING 2005 CALL CENTER VOLUNTEERS

Michelle Achee	Brian King
Leticia Amezcua	Peter Leuck
Marian Apra	Chris Lloyd-Jones
Dale Attias	Jing Luan
Tatiana Bachuretz	Dan Martinez
Claire Biancalana	Rachel Mayo
Noel Burnham	Sandra McCann
Mary Cardenas	Tom McKay
Cathryn Davies	Melissa Molino
Olga Diaz	Tootie Olsen
Glenn Dixon	Manuel Osorio
Georgia Edminster	Margaret Pierce
Terri Evans	Terry Przybylinski
Karen Farrow	Connie Pybrum
Becky Fewel	Elliot Ray
Kim Flock	Lia Reynolds-Bain
Renata Funke	Felix Robles
Gabriel Gutierrez	Ana Ruiz
Laurie Hedin	Belem Ruiz
Julie Herzog	Debbie Soria
Eileen Hill	Ellen Stuck
Mary Beth Hislop	Teresa Thomae
Leah Hlavaty	Carla L. Vaughan
Lyn Hood	Kathie Welch
Sandi Kent	Natasha West

The authors wish to thank the following people for their generous help in preparing for this research study: Erin Lee for managing the day to day Call Center activities; Judy Cassada and Melissa Molina for hand-coding the data from the survey into a database; Eileen Hill for provding training and orientation to Call Center volunteers; and Dr. Dick Borden for providing statistical significance testing.