Computer Science Teachers Association
Analysis of High School Survey Data
(Final Draft)

Eric Roberts and Greg Halopoff May 1, 2005

This report represents the first draft of the analysis of the results of CSTA's 2004 survey of high-school teachers, which generated 1047 responses from teachers in all 50 states and the District of Columbia. This draft is under review by the Research Committee and will presumably be updated as people have more of a chance to look over its findings.

Pre-AP courses (Questions 1 and 2)

The responses to these questions represent something of a surprise. The conventional wisdom has held that the AP curriculum dominates high-school computing courses to the extent that it is, in many jurisdictions, the "only game in town." The survey revealed that many institutions teach computing courses at a lower level than the AP, whether or not they offer the CS AP course itself. The data suggest that approximately twice as many institutions offer a pre-AP course as offer the AP and that the enrollments in pre-AP are on the order of three times larger.

Question 2 provides some interesting insight into the question of what areas are taught in the pre-AP courses, which varies considerably from school to school. The data reported in the summary table show that programming is the most commonly covered topic, followed by hardware, ethics, graphics, and web development, in that order. The summary statistics, however, hide some important factors that CSTA must take into consideration in seeking to understand the state of computing education today.

The first observation that emerges from a more detailed consideration of the data is the high level at which respondents cited "applications" in the "other" category; almost 10 percent of the respondents took the time to include "applications" as a write-in vote, even though it did not appear in the list of options. Given the way people respond to surveys, it is certainly likely that this level represents a significant undercount, in the sense that prompted responses are always favored over free responses. The relatively high showing of "databases" in the survey may offer further evidence in this direction, if some respondents saw that category as encompassing commercial database and spreadsheet applications.

The second important observation is the topic responses are in some cases highly correlated. The matrix in Figure 1 shows the pairwise correlation coefficients for the

Figure 1. Pairwise correlation coefficients for topics in pre-AP course

	app	db	ethic	graph	hdwe	logic	net	prog	sec	web
applications		+0.3	≈0.0	+0.1	≈0.0	−0.1	−0.1	−0.5	≈0.0	≈0.0
databases	+0.3		+0.2	+0.3	+0.2	+0.1	+0.2	−0.3	+0.1	+0.2
ethics	≈0.0	+0.2		+0.2	+0.3	+0.2	+0.2	≈0.0	+0.2	+0.1
graphics	+0.1	+0.3	+0.2		+0.2	+0.1	+0.1	−0.1	+0.1	+0.3
hardware	≈0.0	+0.2	+0.3	+0.2		+0.1	+0.3	−0.1	+0.1	+0.1
logic	−0.1	+0.1	+0.2	+0.1	+0.1		+0.2	+0.1	+0.1	≈0.0
networks	−0.1	+0.2	+0.2	+0.1	+0.3	+0.2		≈0.0	+0.2	+0.3
programming	−0.5	−0.3	≈0.0	−0.1	−0.1	+0.1	≈0.0		≈0.0	≈0.0
security	≈0.0	+0.1	+0.2	+0.1	+0.1	+0.1	+0.2	≈0.0		≈0.0
web	≈0.0	+0.2	+0.1	+0.3	+0.1	≈0.0	+0.3	≈0.0	≈0.0	

specified topics, together with "applications," which was far and away the most commonly cited free response. The largest coefficient, by far, is the –0.5 correlation between "programming" and "applications," which underscores the intuitive expectation that courses that cover applications are less likely to cover programming and vice versa. An analysis of the data clustering would presumably provide more solid evidence, but the strength of this negative correlation (which, for example, is significantly larger than what would seem to be an even more obvious positive correlation between "web development" and "networks") suggests that the sample includes two types of courses:

1. A set of applications-based courses that include relatively little programming. Integrating the responses to Question 7 into the table shows that these courses are more likely (correlation coefficient of +0.1) to be offered under in technology or business departments.
2. A more wide-ranging set of courses that include programming along with a variety of other topics. These courses are more likely (correlation coefficient of +0.2) to be offered under the rubric of science or mathematics departments.

AP courses (Question 4)

As noted in the analysis of the pre-AP offerings, the AP courses reported by the survey respondents tend to be both less common and smaller than some preliminary course in computing. Only 40 percent of the respondents offer the CS AP, and the average reported enrollment is 24. Slightly under half of the respondents indicated that the size of the AP class is between 1 and 10. This level of enrollment underscores a certain fragility in the AP program. If high-school enrollments fall in the way that university enrollments have fallen in recent years, some of these programs, already marginal, seem likely to fail.

It should also be noted that Pre-AP course enrollment is significantly higher, indicating a lack of interest to continue higher level study in the AP course. This finding is consistent with enrollment patterns nationwide as students move from Pre- to AP level courses. This is likely due to the limited number of elective credits students are allowed to take at the high school level (53% enrolled in Pre-AP for elective credit – the largest percentage of responses in Question 1b).

Gender equity (Questions 3 and 5)

The story from the two questions that asked about gender in the pre-AP and AP is reasonably clear. The fraction of women involved in these courses declines from 32 percent in the pre-AP courses to 23 percent in the AP program. This decline is consistent with the many studies that show women falling out of the "academic pipeline" at every stage. Having good data on current numbers in high schools from such a large sample will be helpful to those working in this area at all levels. The numbers are clearly dismally low and create significant barriers to recruitment and retention on the university level.

Trend lines in enrollments (Question 6)

After looking at the data for Question 6, there seems to be no pattern that can provide useful insight into what is clearly one of the central questions. The survey results are split almost evenly between increasing and decreasing enrollments (56 vs. 44 percent). Moreover, that split is mirrored in the state-by-state data. While there is occasionally a large skew in the responses in the state-by-state breakdown, those imbalances occur for small states in which the reporting level is too small to be significant. In all of the states that have a large number of respondents, the numbers are close. The largest number of responses, for example, comes from Texas, where 51 percent cited an increase and 49 percent claimed a decrease in what is clearly a statistical dead heat.

One possible explanation for the seemingly inconsistent reporting of trend lines may be the specification of the time period in the survey instrument. The question asks whether enrollments have "increased or decreased in your school over the past five years." The choice of five years as a time period seems particularly problematic. University enrollments, for example, rose very quickly in the period from 1995 to 2001 along with the dot-com boom, but then declined quite sharply after that. Thus, while it might be clear that enrollments are up from a decade ago and down in the last two years, the question of how they compare to five years ago is far more difficult to answer. Respondents who saw the starting point as being well before the collapse might have focused on the growth, while those who though in terms of the more recent time frame might focus on the decline. Unfortunately, there is no way to determine whether this explanation is in fact the case or whether high school enrollments follow the same patterns that the better-studied ones in universities do. The data produced by this survey are not adequate to draw any such conclusions.

One consistency that can be cited from state to state is the larger percentage of Pre-AP offerings to AP. Two states with the largest number of responses, CA and OH, show a dramatically lower AP offering with 74/44% and 76/35% in Pre-AP/AP, respectively. This trend is relatively consistent among states, and amplifies the disconnect between Pre- and AP computer science. This disparity should be considered as the CSTA makes plans to provide resources to teachers in the CS field.

Certification and computing education (Questions 7 to 10)

Like the issue of enrollment trends discussed in the preceding section, the various questions designed to determine how states differed in terms of their rules for certification provide far less useful information than one would hope. Nationally, the results tend to cluster in the middle, with about the same number of negative and positive responses to each of the yes/no questions. One's intuition would be that this sort of balance masks much more significant diversity at the state level. For example, if half of the states required certification and half did not, the overall numbers would tend to hover around 50 percent without providing any interesting insights. That situation, however, is not supported by the state-by-state breakdowns. The pattern is slightly more skewed than it was for enrollment trends but remains frustratingly inconsistent. The fact that each of the larger states reports a mix on the question of whether a standard curriculum exists could simply reflect large-scale variability by district within states. That view, however, is hard to support in light of the responses to Question 9a, which reads

> Does your state consider CS a certified teachable?

That question poses an issue of fact that should be constant statewide. The responses within most states show a surprisingly inconsistent perception. Nine states, including some with

Figure 2. Greatest challenges

Challenges (decreasing rank)	average rank	#1 fraction
Rapidly changing technology	3.42	31%
Lack of curriculum resources	4.29	13%
Lack of hardware/software resources	4.39	19%
Difficult subject matter	4.46	12%
Lack of student interest/enrollment	4.64	15%
Lack of support/interest by school staff	4.78	13%
Lack of student subject knowledge	5.24	4%
Lack of teacher subject knowledge	5.31	9%

reasonable numbers of respondents like Colorado, split perfectly down the middle on this question, with exactly 50 percent saying that their state considered computer science to be certifiable an the other half taking the opposite view.

The only conclusion that seems to jump out of these data is that the teachers themselves often have a poor understanding about rules and administrative structures within their own state, at least insofar as computer science certification is concerned.

Perceived challenges (Question 11)

Question 11 on the survey asks teachers what they "perceive as the greatest challenges in teaching CS" and then requests a response in the following form:

> Please rank these in numeric order with (1) representing the greatest challenge, (2) the second greatest, and so on:
>
> ___ Lack of hardware/software resources
> ___ Lack of curriculum resources
> ___ Lack of teacher subject knowledge
> ___ Lack of student subject knowledge
> ___ Lack of support/interest by school staff
> ___ Lack of student interest/enrollment
> ___ Difficult subject matter
> ___ Rapidly changing technology
> ___ Other (please specify):

The responses to this request for ranking came out as shown in Figure 2, where they appear in decreasing order of primacy. The table contains two columns, which offer two perspectives. The "average rank" column reflects where this factor came out in the rank orderings provided by respondents. The "#1 fraction" column indicates what percentage of the respondents listed this challenge as the highest priority. Note that the implied orderings are different. Lack of hardware is in second place in terms of the number citing it as the number one concern, but drops to third in average rank. Similarly, even though 15 percent of the respondents identified lack of student interest as their primary concern, it was in the bottom half of the rankings overall. These disparities suggest that the challenges may vary considerably from one institution to another, and that things that are high on some institution's list may not register on another's.

The biggest problem, however, with this set of responses is that a rank ordering gives no indication of absolute importance. Is number 3 on someone's list still important relative to the top two postings, or is it more in the category of a minor annoyance? The data, unfortunately, offer little insight. Some respondents gave an indication along these lines by ranking only a subset of the challenges; about five percent, for example, rated fewer than four of these categories. A handful, moreover, ignored the instructions entirely and gave #1 ratings to more than one challenge, leaving the others blank. In both those categories, "rapidly changing technology" was the most common response, which provides further evidence that people view it as a serious problem.

For this question, looking through the written responses in the "other" category did not reveal any large-scale commonalities. In attempting to categorize those responses, it quickly became clear that the most commonly cited "other" challenge was in many ways a restatement of the "rapidly changing" category applied to something besides the hardware. For example, several people mentioned the language instability in the AP program in the "other" category. Many others, presumably, filed this concern under the "rapidly changing technology" category.

Professional development needs (Questions 12 and 14)

Question 12 asked teachers to rank order their professional development needs. More than with most questions, these responses seem easy to interpret. The "time for training" response was very heavily cited, with an average rating of 1.61 and a #1 ranking level of 58 percent.

The written responses in the "other" category revealed only one significant factor beyond the ones listed: the cost of training and the difficulty of getting those costs reimbursed. This factor was mentioned by 32 respondents, which is about three percent of the total. While this level of response is far below that of the "time for training" category (and presumably interacts with it under the classic "time is money" formulation), the fact that such a large group cited it as a free response indicates that it probably plays a role.

Question 14 then asked teachers to rank what they felt would be the most effective strategies for professional development. Not surprisingly, given the overwhelming importance of "time for training" as a barrier to acquiring such professional development, respondents preferred short workshops to the other forms suggested.

Reasons that students don't choose computer science (Question 13)

This question revealed some very interesting information. The responses to the factors that were suggested are shown in Figure 3, along with the average rank and the percentage of respondents who cited each factor as the number-one concern. Within this group, the competition with other courses for spaces in the academic plan is way out in front, followed by student interest and the perceived difficulty of the subject matter in a close competition for second place. The idea that computer science is perceived to be "geeky" does not rate highly on our respondents' radar screens.

Figure 3. What keeps students out of computer science

Factors (decreasing rank)	average rank	#1 fraction
No room in schedule/academic plan	2.24	43%
Greater interest in other subjects	3.02	17%
Subject matter too difficult	3.09	20%
Elective courses are less “important”	3.26	14%
CS is perceived to be “geeky”	4.07	4%

The problem, however, with this selection of factors is that it leaves out two seemingly significant factors that respondents cited in the “other” category. The largest of these was the lack of availability of courses, either because they have never been offered at that school or because they have recently been dropped. The second reason, running somewhat behind in the number of mentions, is the perception among students, parents, and guidance counselors that there are limited job opportunities in computer science because of the decline of technology in the post-bubble era and the growing use of offshoring. These perceptions run precisely counter to the views of the Department of Labor and most companies, but they clearly have high impact.

As with the other questions for which large groups of respondents chose some “other” category in large numbers, it is difficult to draw any conclusions about the relative ranking between the options that teachers were given and those they had to supply on their own. It seems likely that the competition for time in the schedule would still have come out in first place, but the relative importance of the perceptions about the difficulty of computer science as compared to the perceived lack of opportunity is impossible to gauge.

Demographic data (Questions 15 through 17)

These questions provide data as to the number of years of teaching that the respondents have and the size of the institutions in which they teach. These responses seem entirely straightforward and require little interpretation.

“Other” Data

Further analysis needs to be conducted to categorize “other” responses to questions in the areas of professional development needs, greatest challenges to teaching CS, and departments containing CS.