

PROGRAMMING LANGUAGES AND GENDER

*Comparing differences and similarities in programming language usage
according to programmer gender.*

The scarcity of information technology workers has been widely acknowledged (see [2, 7]). With the downturn in the U.S. economy and the widespread failures of many dot-com companies, the extreme labor shortage appears to be easing somewhat. However, the longstanding technology labor crunch is by no means over. Demand remains for workers with knowledge of specialized technology languages, such as Java and XML programming [5].

A previous *Communications* special section (July 2001) focused on the global IT work force. One of the section's key points was to motivate debate and deliberation regarding the lack of representation of women, minorities, and older workers in the IT work force [1]. The number of qualified IT workers could be increased if additional qualified IT workers could be drawn from these underrepresented groups (particularly women, as they are, by far, the largest of these three groups). The article by von Hellens and Nielsen [11] in the July 2001 section focused on attracting women to the IT field in Australia. In Trauth's article, the focus was on the two work force challenges for Ireland relating to gender and socioeconomic class [8]. Previously, Panteli et al. [6] reported that women were underrepresented in all areas of the IT industry. In an earlier study, Truman and Baroudi [9] found that the mean salary for women IT managers was considerably lower

than males, which they stated, is "a problem suggestive of discriminatory practices."

Our focus here is on a narrower topic: programming languages and gender differences. Specifically, we focus on three questions. First, do female IT workers tend to have different programming languages than male IT workers? Second, if there are differences, are these differences the same for new (less than one year of experience) IT workers as they are for the overall population of IT workers? Third, if there are differences, are they the same for older (40 years old and older) IT workers as they are for the overall population of IT workers?

Overall Survey Results

The results presented here are based on a voluntary Web-based survey on the salaries and languages of information technology workers conducted by dice.com, an online placement company (this survey is available from the

company's Web site, www.dice.com). During the period from June 7, 2000 to April 13, 2001, 37,398 full-time IT workers completed the survey online, providing the data set used in our analysis.

The relevant variables for this article are each respondent's programming languages, years of (total technical) experience, gender, and age. The survey collected data on 49 different programming languages. The average number of languages possessed by males was 3.25 languages, while the average number of languages possessed by females was 2.53 languages. The survey collected data on experience with a six-category question (see Table 1 for a description of the categories). As shown in Table 1, 83.8% of all the respondents were male and 16.2% were female, but the percentage of females increased to 20.6% for the new (less than one year of experience) IT workers. A similar gender split (80.7% male and 19.3% female) occurred for workers with one or two years of experience. For each of the more experienced groups, the gender split was approximately 85% male and 15% female. Hence, the female respondents were, in general, less experienced than the male respondents.

While the percentage of female survey respondents seems quite low, a number of studies indicate the percentage of female IT workers is approximately 20% in many countries [6]. Further, research on gender differences and job searches provides a possible explanation why the percentage of female survey respondents is low—women conduct a job search with less intensity than men [3] and are less likely to separate from an existing job [4]. More important, in our tests for the difference in two proportions for each language, there is no requirement that the two sample sizes be proportional to the two population sizes.

In Table 2a, the results for all the respondents are presented for the 15 most frequently listed ("popular") languages by the respondents. The "Language" column indicates the specific programming language. The "Overall" column denotes the overall percentage that knows this language. The "Male" and "Female" columns denote the percentage of males and females, respectively, who know a particular language. The "Difference" column is the difference between the percentage of males who know a particular language compared with the percentage of females who know that language. The "Significance" column indicates whether the results for the two-tailed test for the difference in two proportions were significant or not, with each letter indicating a greater level of significance (one letter for 0.05, two letters for 0.025, three letters for 0.01, and four let-

ters for 0.001) and the letter indicating whether the results indicated a greater percentage of males or females know this language. (Note: all percentages, including differences, were rounded to one decimal digit.)

In Table 2a, only one difference was negative, indicating that only for Cobol does a higher percentage of females than males know this language. Of the other 14 differences, 12 showed a highly significant language difference for males compared with females. SQL and Oracle both showed a slight, but not statistically significant, language difference of males compared with females.

Paneteli et al. [6] report on several studies indicating the IT culture is a "masculine, engineering type computing culture." In addition, a number of psychological studies indicate that "through experience, people come to share beliefs about the extent

Table 1. Experience and gender.

Experience Level	Experience (in years)	Number of Respondents	Males	Females
1	less than 1	4,176 (11.2%)	3,316 (79.4%) (10.6%)	860 (20.6%) (14.2%)
2	1 or 2	6,140 (16.4%)	4,953 (80.7%) (15.8%)	1,187 (19.3%) (19.6%)
3	3 to 5	12,261 (32.8%)	10,426 (85.0%) (33.3%)	1,835 (15.0%) (30.4%)
4	6 to 10	7,385 (19.7%)	6,315 (85.5%) (20.1%)	1,070 (14.5%) (17.7%)
5	11 to 14	3,009 (8.0%)	2,579 (85.7%) (8.2%)	430 (14.3%) (7.1%)
6	15 or more	4,427 (11.8%)	3,763 (85.0%) (12.0%)	664 (15.0%) (11.0%)
Overall		37,398	31,352 (83.8%)	6,046 (16.2%)

to which tasks are linked to gender" [10]. When viewing this list, one could argue based on these stereotypes that the higher percentage of female Cobol programmers is a result of maintenance on older legacy systems rather than "engineering" new systems. Similarly, the very slight percentage difference favoring males for SQL and Oracle may indicate these languages focus on reporting functions for existing systems rather than engineering new ones.

Examining the data of the less popular languages (not in the top 15 list but still with at least 1% of those surveyed possessing the language), women possessed the following languages in percentage numbers greater than males (in parenthesis: female percentage: male percentage with * indicating the result is significant at the 0.001 level and # indicating the result is significant at only the 0.025 level): CICS (2.8 : 2.0 *), DB2 (4.3 : 3.9), Easytrieve (1.6 : 0.9 *), JCL (4.2 : 3.1 *), Peoplesoft (2.2 : 1.4 *), PL/1 (1.2 : 1.1), PowerBuilder (2.6 : 2.4), SAP (2.1 : 1.6 #), and SAS (1.8 : 1.1 *). Most of these tools

OUR RESULTS SUGGEST WOMEN
HAVE GREATER REPRESENTATION
MAINTAINING COBOL LEGACY SYSTEMS
RATHER THAN ENGINEERING
NEW SYSTEMS USING LANGUAGES SUCH
AS JAVA OR C++.

are used for reporting or analysis functions for existing systems rather than engineering new ones.

Survey Results for New Workers

There is a reason for optimism when focusing on new IT workers (those with less than one year of experience). The percentage of responding female IT workers is 20.6% rather than the overall 16.2%. The difference between the average number of languages possessed by males and females is quite small: 0.35 (2.38 for males and 2.03 for females).

The results in Table 2b show the two most popular languages are HTML and SQL, but the language difference now favors females. Also note that the percentage knowing HTML dropped only slightly from the full sample while the percentage knowing SQL dropped considerably. Surprisingly, Cobol now shows a significant language difference for males compared with females.

Survey Results for Older Workers

Focusing on the older IT workers (40 years old and older), the percentage of female IT workers responding is 19.9%. The difference between the average number of languages possessed by males and females is quite large: 0.79 (2.92 for males and 2.23 for females).

The results in Table 2c show the two most popular languages are again HTML and SQL with the language difference for HTML favoring females (but not significantly). Cobol is the fifth most popular language, and JCL is the twelfth most popular language, with the language difference for JCL favoring females (but not significantly). These results are not surprising given Cobol and JCL's extensive use in older legacy systems.

Language	Overall	Male	Female	Difference	Significance
HTML	35.6%	36.0%	33.5%	2.5%	MMMM
SQL	28.9%	29.1%	28.0%	1.1%	
Visual Basic	23.4%	24.4%	18.3%	6.1%	MMMM
JavaScript	23.0%	23.6%	19.7%	3.9%	MMMM
Java	18.3%	19.2%	13.8%	5.4%	MMMM
C	16.7%	17.8%	10.7%	7.1%	MMMM
C++	16.3%	17.4%	10.6%	6.7%	MMMM
Oracle	13.4%	13.6%	12.7%	0.9%	
ASP	13.0%	13.7%	9.4%	4.4%	MMMM
Perl	11.0%	11.8%	6.7%	5.1%	MMMM
Active X	9.9%	10.7%	5.7%	5.0%	MMMM
Basic	9.4%	9.9%	6.7%	3.2%	MMMM
Visual C++	9.4%	10.2%	5.2%	5.0%	MMMM
OOP	7.6%	8.1%	4.7%	3.5%	MMMM
Cobol	7.4%	7.2%	8.2%	-1.0%	FFF

Significance: 1 letter for 0.05, 2 letters for 0.025, 3 letters for 0.01, and 4 letters for 0.001. Letter: M—a greater number than expected of males possess this language; F—a greater number than expected of females possess this language.

Table 2a. Most popular languages.

Language	Overall	Male	Female	Difference	Significance
HTML	31.1%	30.3%	34.2%	-3.9%	F
SQL	20.1%	20.0%	20.3%	-0.4%	
JavaScript	18.9%	19.3%	17.7%	1.6%	
Visual Basic	18.3%	19.1%	15.2%	3.9%	MMM
Java	17.6%	18.3%	15.0%	3.3%	MM
C++	16.0%	16.8%	12.9%	3.9%	MMM
C	13.2%	13.8%	10.9%	2.9%	M
Oracle	9.7%	9.7%	9.4%	0.3%	
ASP	9.4%	9.7%	8.5%	1.2%	
Basic	7.4%	7.5%	7.1%	0.4%	
Visual C++	7.4%	8.0%	5.2%	2.8%	MMM
Active X	6.2%	6.8%	3.7%	3.1%	MMMM
Perl	6.0%	6.6%	3.8%	2.7%	MMM
OOP	5.0%	5.4%	3.5%	1.9%	MM
Cobol	4.2%	4.4%	3.6%	0.8%	M

Significance: 1 letter for 0.05, 2 letters for 0.025, 3 letters for 0.01, and 4 letters for 0.001. Letter: M—a greater number than expected of males possess this language; F—a greater number than expected of females possess this language.

Table 2b. Most popular languages for new workers.

Language	Overall	Male	Female	Difference	Significance
HTML	24.6%	24.4%	25.5%	-1.1%	
SQL	24.6%	24.6%	24.6%	0.0%	
Visual Basic	20.2%	21.6%	14.7%	6.9%	MMMM
JavaScript	13.9%	14.2%	12.6%	1.6%	
COBOL	13.7%	13.8%	13.4%	0.4%	
C	12.9%	14.1%	7.9%	6.2%	MMMM
Java	12.1%	12.9%	8.8%	4.1%	MMMM
C++	12.1%	13.3%	7.2%	6.1%	MMMM
Oracle	11.3%	11.5%	10.5%	0.9%	
Basic	11.1%	12.0%	7.5%	4.4%	MMMM
Visual C++	7.4%	8.4%	3.3%	5.1%	MMMM
JCL	7.3%	7.1%	7.9%	-0.8%	
Perl	7.2%	7.9%	4.1%	3.8%	MMMM
ASP	7.0%	7.5%	5.3%	2.2%	MMM
Active X	6.9%	7.7%	3.7%	3.9%	MMMM
Perl	6.0%	6.6%	3.8%	2.7%	MMM
OOP	5.0%	5.4%	3.5%	1.9%	MM
Cobol	4.2%	4.4%	3.6%	0.8%	M

Significance: 1 letter for 0.05, 2 letters for 0.025, 3 letters for 0.01, and 4 letters for 0.001. Letter: M—a greater number than expected of males possess this language; F—a greater number than expected of females possess this language.

Table 2c. Most popular languages for older workers.

Conclusion

While it is very important to draw more women to the IT field, it is equally important to ensure these new workers are welcomed fully into the IT community. Our results suggest women have greater representation maintaining Cobol legacy systems rather than engineering new systems using languages such as Java or C++. In a similar vein, the overall slight percentage difference favoring males (compared to the huge percentage difference favoring males for Java or C++) for SQL and Oracle may indicate that women focus more on reporting functions for existing systems rather than engineering new ones. In terms of new IT workers, a significantly greater proportion of females indicates knowing HTML compared to their male counterparts. While HTML may be an entry point to an IT job, it may also stereotype an IT worker into less demanding programming jobs.

Paneteli et al. [6] recommends a change in "the attitudes of employers, male colleagues, and managers. Male-dominating attitudes and perpetuating stereotypes seem to predetermine the positions that should be held by women ... Indeed, women themselves may come to believe the stereotypes." ■

REFERENCES

1. Arnold, D. and Niederman, F. The global IT work force. *Commun. ACM* 44, 7 (Jul. 2001), 31–33.
2. Ferratt, T., Agarwal, R., Moore, J., and Brown, C. Observations from 'The Front': IT executives on practices to recruit and retain information technology professionals. In *Proceedings of SIGCPR 99* (1999), 102–112.
3. Keith, K. and McWilliams, A. The return to mobility and job search by gender. *Industrial and Labor Relations Review* 52, 3 (Apr. 1999), 460–477.
4. Kulik, L. A comparative analysis of job search intensity, attitudes toward unemployment, and related responses. *Journal of Occupational and Organizational Psychology* 73 (Dec. 2000), 487–500.
5. Mortl, J. Labor shortage eases. *Internet Week* (May 8, 2001).
6. Panteli, A., Stack, J., and Ramsay, H. Gender and professional ethics in the IT industry. *Journal of Business Ethics* 22 (1999), 51–61.
7. Stewart, T. In search of elusive tech workers. *Fortune* 137, 3 (1998).
8. Trauth, E. Mapping information-sector work to the work force. *Commun. ACM* 44, 7 (Jul. 2001), 74–75.
9. Truman, G. and Baroudi, J. Gender differences in the information systems managerial ranks: An assessment of potential discriminatory practices. *Management Information Systems Quarterly* 18, 2 (Jun. 1994), 129–141.
10. Vancouver, J.B. and Ilgen, D.R. Effects of interpersonal orientation and the sex-type of the task on choosing to work alone or in groups. *Journal of Applied Psychology* 74, 6 (1989), 927–934.
11. von Hellens, L. and Nielsen, S. Australian Women in IT. *Commun. ACM* 44, 7 (Jul. 2001), 46–52.

RONALD DATTERO (rsd686f@smsu.edu) is an associate professor in the Department of Computer Information Systems at Southwest Missouri State University.

STUART D. GALUP (sgalup@fau.edu) is an assistant professor in the Department of Information Technology and Operations Management at Florida Atlantic University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

© 2004 ACM 0002-0782/04/0100 \$5.00

Stay on Top of ACM News with MemberNet

Coming in future issues of MemberNet:

- A new look
- Findings from our MemberNet reader survey
- New computing initiatives to reach underserved groups
- And much more!

All online in MemberNet: www.acm.org/membernet.