





Figure 1. Time on task.

growth. We have not analyzed what percent of the communication is on task (about programming or writing), versus being purely social. Furthermore, even when communication is purely social, it is not clear to what extent this contributes to the development of writing skills. Thus, this finding is intriguing but difficult to interpret.

## 5. SCORING OF STUDENT ACHIEVEMENT

In 2000, we used portfolio scoring techniques to analyze students' programming achievement on MOOSE Crossing according to the following scale:

- 0: Wrote no scripts
- 1: Demonstrated understanding of basic input/output
- 2: Used variables and properties
- 3: Performed list manipulation and flow control
- 4: Demonstrated mastery of all aspects of the system

These ratings were produced by two human raters. In cases where the raters disagreed, a third person rated the student's level of accomplishment. This technique was applied to a random sample of 50 participants (Bruckman et al., 2000). Subsequently, we became concerned that perhaps our categories















as a form of writing. For at least this one child (and presumably some others), this environment has made computer programming more appealing than it likely otherwise would be. But if differences in achievement persist even in this environment, what is their source? Data analysis presented in this paper suggest that educators wishing to increase girls' level of technical achievement should explore strategies for increasing girls' interest in technical subjects. In both the BASIC programming environment Marcia Linn studied in the early to mid-1980s and in the CSCL environment we designed and studied from the mid-1990s to the present, girls program equally well as boys when they devote equal time to the activity.

## 9. CONCLUSION

A good log file is a terrible thing to waste. Use of log file data as a research tool can benefit from:

- Complementary use of qualitative and quantitative methods
- Both manual and automated methods of analysis
- Strategic leveraging of regularities found in the data from a particular domain
- Recognition of how 'human readable' a particular data set is, and how this shapes the kinds of analysis possible
- Use of computerized tools to aid the human reader
- Complementary use of data from other sources (such as interviews, field notes, and surveys)
- Attention to ethical issues concerning the recording and analysis of log file data, maintaining respect for individuals' privacy and their rights as human subjects (in situations in which the work constitutes 'human subjects research')

Effectively leveraging these strategies, log files can provide a wealth of information about online communities and the learning taking place there.

## ACKNOWLEDGMENTS

Carlos Jensen and Austina DeBonte contributed to earlier versions of this chapter. The authors would like to thank everyone who has contributed to the MOOSE Crossing project over the years, especially Mitchel Resnick, Alisa Bandlow, Elizabeth Edwards, Naureen Hooda, Justin Jackson, Albert Lin, Tysen Perszyk, Will Scott, Adam Skwersky, Trevor Stricker, Steven Tamm, and Joseph Uhl. Jason Elliott and Jon Orwant provided help with Perl scripts. Christian Jensen assisted with statistics. Special thanks to the kids, teachers, and parents who have participated in MOOSE Crossing. Research in the















