

Technical Report N° 2000/04

***Data collection techniques used to study programming -
and AESOP***

Carina Paine

***Department of Computing
Faculty of Mathematics and Computing
The Open University
Walton Hall,
Milton Keynes
MK7 6AA
United Kingdom***

<http://computing.open.ac.uk>

|

Data collection techniques used to study programming - and AESOP

Carina Paine

1. Introduction

All empirical studies of programmers require suitable methodologies and techniques. Different methods of data collection are appropriate in different circumstances. Surprisingly, the initial research in the study of programming focused on experiments, when observational methods would seem more appropriate for a new discipline [1]. Recently there has been a shift to observational methods due to their ability to capture more complex aspects of programming. "I don't think that programming is too complex a behaviour to be studied, but it may be too complex to be studied in a laboratory" [2].

AESOP, *An Electronic Student Observatory Project*, is a collection of computer-based data collection tools. Our educational environment is one in which students study independently at a distance, off-line, using software developed for the Open University course M206, *Computing: An Object-oriented approach*. The aim of the project is to discover how students learn to program, by observing students unobtrusively, electronically, and automatically, and to record these observations in a manner that is useful for both instruction and research.

This document reviews the differences between experimental and non-experimental methods, and discusses the advantages and disadvantages of the various data gathering techniques in the empirical study of programming, including new computer aided techniques. The data collection techniques used for AESOP are also described, and the reasons for our choice of techniques is explained.

2. Methods

Among the variety of research methods for collecting data there is a divide between non-experimental methods and experimental methods. The main differences between these methods are described below.

2.1 Non -Experimental Methods

Description - A form of exploratory data collection.

Use - When the research theme is novel and there is a lack of information to engage in the other types of data collection
When existing theoretical frameworks are unable to precisely define hypotheses.

Advantages

The results can be generalised

Disadvantages

There is no manipulation of variables

It can be difficult to collect and record data

The resultant data can be unclear, and undefined

There is a lack of control over extraneous variables

2.2 Experimental Methods

Description - Involves hypotheses testing and controlling / manipulating variables

Use - When there is an existing theoretical framework and defined hypothesis.

Advantages

There is a decreased "cost" of bulk data from observation

Produces well defined information

Disadvantages

Inability to generalise (from laboratory experiments)

Subjects' behaviour may be constrained by an experimental situation

Subjects' are separated from their normal working environment

An artificial environment is not representative of real life use of a system

3. Data Gathering Techniques

There are five basic data gathering techniques

3.1 Observation

Description - Data gathering where researcher makes notes of subjects' behaviour

Use - A method of monitoring subjects

Advantages

Record of what subjects actually do rather than what they think they do

Generalisation of results if subjects are in normal environment

Disadvantages

Can be intrusive which may influence subjects' performance (Hawthorne effect [3])

Manual data capture is not efficient. Humans are easily distracted / bored, and have difficulty coping with recording sudden peaks of activity

Inference of the observer can affect the findings

Very time consuming

Cause and effect relationships cannot be easily demonstrated

3.2 Questionnaire

Description - Students' study a program and then are asked questions about it. Can be used alone or with other protocols

Use - To determine comprehension of a program

Advantages

Simple technique

Produces large databases of information quickly and economically

Closed questions are easy to analyse

Open questions are easy to construct and provide independent opinions

E-mail / web questionnaires provide real time tools for collecting responses automatically as well as analysing and creating representations of results

Disadvantages

Closed questions are difficult to construct as it is hard to anticipate what Ss may want to say and interesting / unexpected answers may remain unknown

Only useful for evaluation attitudes, not able to provide objective data

3.3 Interview

Description - Face to Face questioning. Can be unstructured or structured

Use - To obtain subjects answers to questions

Advantages

Structured interviews are more systematic and can provide more coverage of a topic

No questions will be missed by mistake

New questions can be added if something interesting arises out of previous responses

Controlled data gathering assists validity (for example, the validity of answers can be checked by repeating the question)

Disadvantages

If self-administered (by the researcher) can be biased

High cost

Time consuming

3.4 Think / Talk Aloud

Description - Subjects asked to vocalise their thoughts this can be while performing a programming task, or after the sequence of events

Use - To investigate human problem solving behaviour (protocol provides direct evidence for cognitive processes while performing complex tasks)

Advantages

What subjects are thinking rather than trying to infer indirectly from actions

A rich source of information

Disadvantages

Disrupt primary task of interaction with system, and interfere with subjects' normal programming style through increased workload

Individuals are different in their capacity to articulate their knowledge. Some people think faster than they talk

Subjects may work automatically and fail to verbalise routine tasks

Problems of interpretation; protocols are not used in raw form, they are coded.

Problems with consistency and accuracy among coders Training can increase reliability and consistency among coders but still is a subjective process prone to differences

Difficult to get reliability if coding a long protocol

Carrying out analysis is difficult / time consuming / expensive, and as a consequence almost always uses a small number of subjects

3.5 Video Recording

Description - Programming session videotaped with camera and microphones to observe programmer's behaviour

Use - Provides additional information by showing what a subject is doing when they are not thinking aloud or typing on the computer keyboard, and what is being displayed on the computer screen.

Advantages

Video Equipment allows researcher to record the programming situation without the presence of a cameraman

Can capture the full dynamic nature of natural programming activity

Disadvantages

Large volumes of data; which involves time consuming analysis and high costs

Intrusive; subjects may find it difficult to relax in front of the camera, which in turn may influence subjects' performance

4. Computer aided data gathering techniques

Ideally unobtrusive techniques are needed for programming studies, where subjects are unable to see or hear the data gathering equipment. A range of techniques which are becoming increasingly sophisticated are those which make use of computer aided tools.

4.1 Keystroke

Description - Data comprised of each key pressed by a subject while working with a computer system

Advantages

Data may be recorded passively and unobtrusively

Disadvantages

Large volumes of fine-grained data, which can lead to problems in storing / managing / interpreting

Inferring what keystrokes add up to in terms of the commands and functions supported by the application program requires an analytical program

Some systems may miss many user actions

Problem of completeness from systems that use graphic displays and mouse control, many user actions will not be recorded

Only records the input stream

4.2 Data logging / capture

Description - A form of observation (unique to computer studies)

(Often combined with video recording)

Advantages

Passive and unobtrusive

Direct reading of user behaviour by computer

Data capture can help to clarify actual programmer activity

Provides detail over a long period of time

Disadvantages

Large volumes of data, which can lead to problems in storing / managing / interpreting

Only direct actions on the system are monitored, no additional information is recorded to show other subject activities

4.3 Action Level Protocols [4]

Description - Records protocols at the level of user actions (rather than keystrokes).

Advantages

Passive and unobtrusive

Captures data missed by keystroke protocols

Recorded in machine-readable form eliminating need for manual transcription

Disadvantages

Problems of analysis and interpretation remain

4.4 Visus System [5]

Description - Continuously records the contents of text displays and collects it into the disk file.

Advantages

Passive and unobtrusive

Totally based on software, is cheap and easy to use.

Disadvantages

Problem of completeness from systems that use graphic displays and mouse control, many user actions will not be recorded

5. AESOP

The AESOP project is interested in exploratory data collection, and uses a non-experimental method, involving electronic observation and questionnaire techniques for the study of student programmers. It is important to have a setting that can elicit the behaviours that we want to observe. Unlike an experimental setting our students are in their natural working environment. AESOP uses real programming tasks to observe real programming activity, which enables our results to be generalised.

Unlike the other data gathering techniques described above, AESOP's electronic data capture is invisible to the student and does not interfere with the students' work. Therefore, AESOP is less likely to influence students' performance. Even if students are initially aware of being recorded it has been shown that stress diminishes with time [6].

It can be costly, both in time and money to collect data from a large number of subjects using the described data collection techniques, AESOP's data capture is automatic and records all significant events¹. Large volumes of data can also mean that analysis is difficult, time consuming and expensive, and as a consequence a lot of studies use a small number of subjects. We want mass observation, and so apart from the fact that our data capture is automatic, we have also built a number of automated tools to assist with analysis [7].

The main difficulties of manual data capture have been overcome by computer aided data collection. AESOP's electronic observation is more advanced than keystroke techniques as it captures both text and graphics displays. Input *and* output streams are recorded by AESOP, showing what is being displayed on the students' computer screen. The system is cheap and easy to use as the students system is modified only very slightly (the recorder software is provided to students as two small files). The recorded data is in the final form of a text file with textual descriptions of recorded events for readability (by machine and by humans) eliminating the need for manual transcription.

As mentioned above one of the disadvantages of observation techniques is that cause and effect relationships cannot be easily demonstrated. We cannot tell what caused a particular event to occur. We have to go through a sequence of events and we do not know the cause of each event. Although the inference of the observer is taken out by the fact that data capture is electronic and automatic, inferences may be made when replaying the recording.

¹ Events that we decided were important, e.g. page turns, button clicks, hyperlink selection and expression evaluations.

Although M206 is highly constrained there is still a lack of control over extraneous variables. We do not know what a subject is doing when they not working on the computer. The use of video recording or talk aloud techniques may provide us with this valuable information.

Although data capture is automatic we are reliant on students sending the recordings back.

Questionnaires are part of AESOP methodology, but are separate form the recorder. They are only used to supplement observations, to obtain data about previous computing experience, and attitudes towards our research. The questionnaires use both open and closed questions. The use of an online questionnaire means that no questions will be missed by mistake by students, as they cannot submit their questionnaire unless they have answered all of the questions.

6. Conclusions

The advantages and disadvantages of AESOP in alternative data collection techniques are highlighted above. The study of computer programming requires a wide range of approaches to the collection of data, and ideally the techniques whose advantages AESOP does not meet could be used as supplements to the recorder for data collection.

The talk aloud technique identifies what subjects are thinking, rather than trying to infer indirectly from actions. At present we can only guess what subjects might be thinking. This technique would be especially useful to us to determine the reasoning behind a subject's actions when an error occurs.

The video recording technique captures the full dynamic nature of natural programming activity, by showing what a subject is doing when they are not thinking aloud or typing on the computer keyboard. This technique would be useful to us to determine what subjects are doing when we observe breaks in activity in our recordings. However, these techniques do have their own disadvantages, and one of the most important features of AESOP is that it is an unobtrusive system.

References

- [1] Gilmore, D. J. (1990) Methodological Issues in the Study of Programming. In Hoc, J.M., Green, T., Samurcay, R. Gilmore, D. (Eds) *Psychology of Programming*. London: Academic Press.
- [2] Weinberg, G. (1999) *The Psychology of Computer Programming, Silver Anniversary Edition*. New York: Dorset House Publishing.
- [3] Roethlisberger, F. & Dickson, W. (1939) *Management and the Worker - An Account of a Research Program Conducted by the Western Electric Company, Hawthorne Works, Chicago*. Cambridge: Harvard University Press.
- [4] Smith, J.B., Smith, D.K. & Kupstas, E. (1991) *Automated Protocol Analysis: Tools and Methodology*. Technical Report 91 – 034. University of North Carolina, Chapel Hill.
- [5] Kivi, M.R., Gronfors, T. & Koponen, A. (1998) MOTHER: System for continuous capturing of display stream. *Behaviour and Information Technology*, 17 (3), 152 - 154.
- [6] Timpka, B. E. & Arborelius, E. (1991) A Method for Study of Dilemmas during Health Care Consultations. *Medical Inform.*, 16 (1).
- [7] Macgregor, M., Thomas, P.G. & Woodman, M (1999) *Recording User Actions in a Smalltalk Programming Environment*. TOOLS USA.

Bibliography

Brooks, R. E. (1980) Studying Programmer Behaviour Experimentally: The problems of Proper Methodology. *Communications of the ACM*, 23, 207 - 213.

Coolican, H. (1990) *Research Methods and Statistics in Psychology*. London: Hodder and Stoughton

Downton, A.C. (1991) Evaluation techniques for human – computer systems design. In Downton, A.C., (editor) *Engineering the human – computer interface*. London: McGraw-Hill.

Kivi, M. (1999) *Empirical Methods used to Study Programming: A Computer-Aided Approach*. Licentiate Thesis. University of Kuopio

Kivi, M.R., Gronfors, T., Koponen, A. (1998) MOTHER: System for continuous capturing of display stream. *Behaviour and Information Technology*, 17 (3), 152 - 154.

Shneiderman, B. (1980) Software Psychology. *Human Factors in Computer and Information Systems*. Cambridge: Winthrop Publishers, Inc.

Shneiderman, B. (1986) Empirical Studies of Programmers: The territory, Paths, and Destinations. In Soloway, E., Iyengar, S. (Eds) (1987) *Empirical Studies of Programmers*. New Jersey: Ablex Publishing Company.

Sheil, B. (1981) The Psychological Study of Programming. *Computing Surveys*, 13, 101 – 120

Smith, J.B., Smith, D.K., Kupstas, E. (1991) *Automated Protocol Analysis: Tools and Methodology*. Technical Report 91 – 034. University of North Carolina, Chapel Hill.

Weinberg, G. (1971) *The Psychology of Computer Programming*. New York: Dorset House Publishing

Weinberg, G. (1999) *The Psychology of Computer Programming, Silver Anniversary Edition*. New York: Dorset House Publishing.