

An Electronic Clipboard for Field Data Collection

Todd C. Wehner

Department of Horticultural Science, North Carolina State University, Raleigh, NC 27695-7609

Horticultural researchers are often in the position of having to collect a lot of data for their experiments, and then get it analyzed and summarized in a short period of time, and all that within the constraints of a limited budget. I have converted to a new data collection system recently which offers a number of advantages over the old system of writing the data on sheets of paper, entering the data from the data sheets into the computer for analysis, and then writing the article using a "rough" printout from the computer (Fig. 1). This is a good time to convert to electronic data collection because the necessary equipment is easy to use and within reach of many research budgets. Computer programming skills are not required to set up and use the system, since all of the programs used are available commercially, and are easy to learn.

The general procedure is to collect data on a portable microcomputer using a word processing program, then transmit the dataset to a desktop microcomputer for printing, and storage on disks. Subsequently, the data may be analyzed using either a statistical program on the microcomputer, or by transmitting it to a mainframe computer for analysis using programs such as the Statistical Analysis System (SAS Institute, Cary, North Carolina).

Equipment. The data collection system that I use consists of a Tandy 200 portable microcomputer and an Apple Macintosh Plus desktop microcomputer. The Model 100 portable microcomputer is slightly smaller and lighter, and also less expensive, but the Tandy 200 has a larger screen and better keys for cursor control (used frequently during data collection). Accessories useful to have are an external, rechargeable battery for the Tandy 200 (to increase the time that one can go without replacing or recharging the internal AA size batteries), and 2 extra banks of random access memory for a total of 72 kilobytes. The Macintosh system includes a second disk drive, an Imagewriter II printer, MacWrite version 2.2 (the new 4.5 version is not as easy to use for documents that are handled as text only) word processing program, and Red Ryder 7.0 communications program.

The procedure for data collection is similar to the old method, but is faster and easier since the data is not transcribed from paper to computer, saving additional time spent previously in proofreading.

Procedure. The system we used until 1984 involved paper data sheets which were held in binders until needed (Fig. 1). Data sheets were transferred to a clipboard for data collection, then photocopied and given to data entry services to type into a file on the mainframe computer used for data analysis. The data was then printed for checking against the original and mistakes corrected. Finally, the data was printed, analyzed using a SAS program, and the output used to write a report.

The new system involving electronic data collection was tested in 1983 and 1984 with 2 data collectors (a Datamyte 1000 and a Tandy 100) to evaluate the performance under field conditions which included exposure to rain, dust

and physical abuse. In 1985, the present system was used exclusively with no backup data sheets since we felt that the problems with the system had been mostly solved.

The new system involves writing blank data sheets on the desktop computer and transferring the data sheets to the portable computer the day before needed. After collecting the data, it was transferred back to the desktop computer, and stored and printed. When all of the data was collected for an experiment, the data was combined with a SAS program and sent to the mainframe computer for analysis. The results were then transferred from mainframe to desktop computer in a form that was nearly ready for publication.

Example session. The portable microcomputer used for data collection (the Tandy 200) has a screen that is 40 columns wide. Therefore, the data sheet is limited to 39 columns of text (leaving the last column in each line for the carriage return). That is usually sufficient for the data we collect during one session.

An experiment might involve collecting yield data for 6 harvests, quality data for 3 samples, and disease data at the end of the season. For such an experiment, we would use 8 data sheets: 6 for yield, and 1 each for quality and disease. The yield and quality data are collected simultaneously during 3 of the harvests using 2 microcomputers (previously, we used 2 clipboards). Thus, even for an experiment involving much data, the 40 column screen is adequate. The data is collected at different times or in different places, and is, therefore, easily divided into sets. An example of the way one of our data sheets looks is shown in Fig. 2.

Problems. The problems we have noticed with this system are not serious, and are easily solved. Battery power is a limitation in the Tandy 200, since the internal AA alkaline cells last about 14 hr. For that reason, we generally use an external rechargeable battery as mentioned above. The electrical contact of the "down arrow" cursor key of the Model 100 became worn after one season of data collection (since it was used to go down to the next line of data, each line corresponding to a plot number). For that reason, we started using a Tandy 200, which has better cursor keys. Dust in the computer was not a problem, but rain was. Both of those problems can be solved by enclosing the computer in a clear plastic bag and typing on the keys through the bag.

With the old system, it was easy to train new workers to enter data on paper. The new system required additional time to teach the operation of the Tandy 200. However, that was generally not a problem for us since we use the same 2 or 3 people to collect data all year. Another problem we noticed in one experiment is that it is difficult to collect data on a portable computer that is not in order. When data is being called out by many people for one person who is writing in several different places on the data sheet, it is easier to use pencil and paper. The experiment can usually be planned to avoid such confusion, however.

Finally, the communication program in the Tandy 200 does not add a line feed to the carriage return at the end of each line when transmitting to another computer. Therefore, when unloading data from the portable, it is necessary to have a communications program in the desktop computer that will add a line feed. The program, Red Ryder, for the Macintosh will do that.

Advantages. A number of advantages of the electronic data collection system were apparent after 2 seasons of field use. It is easier to use when working under windy or rainy conditions (which formerly made it difficult to handle paper data sheets). Once the data has been collected, a report can be generated within hours with little additional effort. That helps to cut labor costs, and makes it possible to base field selection in a breeding program on complete data (for example, summarized over replications and locations).

Figure 1. Procedures for data collection using paper data sheets, or electronic datasets on a portable microcomputer.

Paper	Electronic
Write out data sheets on paper	Make up datasets on desktop computer and print on paper
Store data sheets for each experiment in binder	Store datasets on disk and put printout in a binder
Move data sheets for present experiment to a clipboard	Transfer datasets for present experiment to portable computer
Fill in the data sheet	Fill in the dataset
Photocopy the data sheets when finished	Transfer dataset to desktop computer
Give photocopy to data entry service	Store on disk and print on paper
Print dataset when entered on computer	
Proof photocopy against computer printout	
Correct errors on computer dataset and print	
Write program for data analysis	Write program for data analysis
Analyze the data and print	Transfer data with program to mainframe computer for analysis
Write report	Transfer analysis summary to desktop computer and write report

Figure 2. Example of a 40-column data sheet as it would appear on the Tandy 200 screen (as well as on the computer it was transferred to)².

Genotype x Environment Study
 Summer, 1985

DATES

Planted-

Clayton: 7/12
 Clinton: 7/18
 Stress: 7/16
 Castle Hayne: 7/17

Harvested-

Clayton: 8/28 and 9/11
 Clinton: 8/29 and 9/5
 Stress: 8/29 and 9/5
 Castle Hayne: 8/30 and 9/5

Identification													Quality			
-----													-----			
													Fruit No			
Yr	Lc	nr	Plot No.	R	Ln	St	No	Ct	To	Cl	OS	h	S	C	S	O
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
85	CA	M	1	0353	1	19	30	36	03	05	5	5	3	6		
85	CA	M	1	0354	1	06	30	29	05	05	2	5	3	6		
85	CA	M	1	0355	1	01	30	30	03	05	3	8	6	8		
85	CA	M	1	0356	1	16	30	22	03	05	3	7	4	7		
etc																
85	CH	F	2	1056	2	01	30	08	04	00	8	8	6	7		

²The identification columns were filled out when the experiment was planned. The dates and the remaining data columns were filled out during the data collection operation. Column headings are abbreviated to help the person collecting the data to remember what was needed. The headings are: Year, Location, Season, Crop, Plot Number, Replication, Line Number, Stand Count, Total, Cull and Oversized Fruit Numbers, and Shape, Color, Seedcell and Overall Performance Quality Score.