

METHODS & DESIGNS

BORES: Behavior observation recording and editing system

EUCLID O. SMITH

*Yerkes Regional Primate Research Center, Lawrenceville, Georgia 30245
and Department of Anthropology, Emory University, Atlanta, Georgia 30322*

and

MICHAEL L. BEGEMAN

Austin, Texas 78703

This paper describes the use of a Datamyte 900 solid state data collection system for observational research. Data editing and reformatting programs that organize the behavioral data into a consistent format for use with widely available statistical packages and edit the data for logical inconsistencies are discussed. Data collection using both the focal animal observation and group scanning, or scan-sample, techniques are amenable to this approach.

The use of a solid state digital data collection device has many advantages over traditional social behavioral data collection techniques used in observational research (see Holm, 1978, for a review of various data collection techniques and their associated strengths and weaknesses). In this paper we describe the use of a Datamyte 900 system, its interface with a minicomputer, and subsequent editing and formatting of behavioral data.

Several different types of solid state keyboard computer-compatible data collection systems have been described (e.g., Celhoffer, Boukydis, Minde, & Muir, 1977; Fitzpatrick, 1977; Stephenson & Roberts, 1977), but we decided on the Electro/General Datamyte Model 900 to satisfy our particular needs of portability, ease and flexibility of operation, and reliability. The Datamyte has been used for various types of human and animal behavioral observations (for various uses of Datamyte systems in human social observational researches, see Sidowski, 1977, pp. 403-455). Although the data described in this paper were collected on monkeys, the techniques are applicable to similar animal and human social activities.

DATAMYTE 900

The Datamyte 900 is a battery-powered micro-processor-based device for the collection, storage, and interfacing of data. Data are entered via one of the five input modes on a 19-key alphanumeric keyboard with

This research was supported by NIDA Grant DA-02128 and NIH Grants RR-00165 and RR-00167.

10 numerical (0-9), 4 alpha (C, F, H, and *), and 5 control keys, and they are stored in the internal memory of the Datamyte. The user can specify one of four possible input modes for collecting different types of data (In-Mode 1, data and time; In-Mode 2, one-character autotime; In-Mode 3, two-character autotime; In-Mode 4, data without time). For ease of data manipulation and transmission, the user can also select from eight different output modes that format data with or without line numbers, in columns or in lines, with carriage return and line feed appended, or with carriage return only appended. The Datamyte has other additional features: an internal interval timer for producing auditory tones at any 1 of 10 preselected intervals, a remote input receptacle for electronically entering keyboard functions, a choice of six different transmission rates, and an internal standby battery that allows the contents of memory to be saved while rechargeable nickel-cadmium battery packs are changed. Also, the unit can be operated while connected to a conventional ac power source. To insure maximum flexibility in designing research protocols, each unit was configured with 32K bytes RAM, the maximum currently available.

BORES: BEHAVIOR OBSERVATION RECORDING AND EDITING SYSTEM

BORES was designed for collecting social behavioral data on a group of stump-tail macaques (*Macaca arctoides*) living in a 30.48 x 30.48 m open-air compound at the Yerkes Regional Primate Research Center Field Station near Lawrenceville, Georgia. The specific aim of the project is to evaluate the social behavioral effects of

acute administration of certain abused drugs on individual group members and the reaction of the group to the experimental subjects. Our general strategy is to collect data on the daily activities of individual animals living in a social group, as well as a profile of the activities of the group as a whole. Two different research protocols are involved. After the data are collected, they are transmitted in ASCII code to a PDP-11/03.¹ One of the features of the Datamyte is that there is no intermediate step necessary for data transmission. Simple connecting cables interface the Datamyte with a variety of peripheral devices. After the data have been transmitted and stored on magnetic disk, two data editing programs are executed. These interactive editing programs check for correct syntax and logical inconsistencies. After editing, the data are stored in disk files in a standardized format suitable for further analysis, and they are cataloged by subject and date.

Observational Techniques

Focal animal technique. Observational data are collected using both the focal and group scan techniques (see Altmann, 1974, for a review). The focal animal observational technique involves observation of individual animals for predefined time intervals² and basically answers the questions of who does what to whom, when, and in what sequence. Data are entered into the Datamyte using the data plus time entry mode (In-Mode 1). Each observation session is preceded by a set of information that allows us to partition sequential focal animal sessions, as well as to maintain a record of certain environmental conditions for the observation session. Using the alphanumeric capabilities of the Datamyte, we enter five sets of header information: (1) date, (2) observer number, (3) weather conditions, including information about cloud cover, wind conditions, and temperature,³ (4) focal animal number, and (5) time (24-h clock). Figure 1 provides an example of focal animal header information.

After the header information is keyed in and the observer is ready to begin, the internal clock on the Datamyte is started. Since one animal is identified as focal for each observation session, the focal animal is implied in our data collection scheme. Thus, if the focal animal initiates a social interaction, we key-in only the type of interaction/recipient. On the other hand, if another animal initiates an interaction with the focal animal, the initiator/type interaction is keyed (entered). In the former case, the data string is 012*27 (ENTER); in the latter case, the data string 27*012 (ENTER) appears. The 012 signifies the behavior code, 27 is the interactant's code, and * is a field separator. Depressing the ENTER key not only enters the data into the Datamyte memory, but also records the time of that entry with respect to the start of the observation session in hundredths of a minute. This is extremely important in analyses that are socially significant for both fre-

FOCAL HEADER

H1092479	DATE
H202	OBSERVER NUMBER
H31165	ENVIRONMENTAL CONDITIONS
H4001	FOCAL ANIMAL CODE
H50800	TIME

GROUP SCAN HEADER

H1092479	DATE
H202	OBSERVER NUMBER
H31165	ENVIRONMENTAL CONDITIONS
H416	GROUP SIZE
H50800	TIME

Figure 1. Examples of header information for focal and group scan observation session.

quency and duration of occurrence. For example, the behavioral event "hit" is significant in terms of its frequency of occurrence per unit time, while "grooming" is significant for its duration as well. The statement that two animals groomed 12 times in an observation session tells us something entirely different from the statement that two animals groomed 2 times for an average duration of 3 min. To accurately record durations of behaviors, we must specify a beginning and a termination point. This is done through the use of both explicit and implicit terminators. An explicit terminator for the behavior "groom" is obviously "terminate groom." The groom and the terminate groom then delimit the grooming bout. Given the associated times with these events, it is easy to determine the duration of the grooming bout. For active, highly social animals such as macaques (or humans), the observer can forget to key in a termination code for a durational behavior. To overcome this, a subroutine identifies potential implicit termination behaviors and inserts the proper explicit termination code into the data stream if the given criteria are satisfied. For example, an implicit terminator for a "groom" is a "chase." (See the section on data editing for details of this subroutine.)

The timing feature of the Datamyte is also important in assessing interobserver reliability. By having time associated with each data entry, it is possible to accurately assess interobserver reliability using, for example Cohen's (1960) Kappa, a statistic that measures nominal scale agreement between two sets of data (see Hollenbeck, 1978, for a review of the problems of interobserver reliability).

Figure 2 shows an example of the raw data from three focal animal observation sessions. Each focal

data set is separated from another test by a series of header information. The clock is shown as 00000. Thereafter, following each interaction is the associated time. In the first set of data, the first interaction occurs .21 min after the start of the observation session.

Group Scanning Technique

Group scanning or scan-sampling techniques provide different types of data from focal observations. Scan sampling involves recording an individual's current activity at preselected moments in time (e.g., 1-min intervals). Data are obtained on the time distribution of behavioral states for the entire social group and provide data that are almost impossible to obtain using other techniques (Altmann, 1974). In observations of group-living stump-tail macaques (*Macaca arctoides*), we are interested in sampling the interactions of selected individuals and assessing the activity patterns of the entire group.

For group scan observations, the simple data entry method is used (In-Mode 4). All animals in the group are surveyed at 1-min intervals, and their behavior is classified into one of seven categories: (1) aggression, (2) submission, (3) general social, (4) affiliative, (5) play, (6) sexual, and (7) self-directed or solitary. The header information for the scan samples is similar to that entered for the focal animal sessions, except that instead of a focal animal number being entered as Item 4, the group size is entered. Figure 1 provides an example of header information for group scan data.

Figure 3 is an example of the raw data generated from three group scan sessions (0800, 1000, 1400 h). Each string of digits between the * delimiters represents one group scan sample. The digits indicate the behavior class performed by each animal in the group as it was surveyed. Variable-length records, extending to multiple lines, if necessary, are permitted. Further, no times are associated with these data, as they are collected at predetermined intervals.

```
H1092479
H202
H31170
H4001
H50800,00000,
001#19,00021,19#030,00024,045#19,00027,040#05,00033,041#05,00036
05#052,00039,128#05,00045,05#043,00047,31#086,00051,31#137,00054
045#31,00058,101#06,00066,
H1092479
H202
H31170
H4001
H50815,00000,
001#19,00007,19#030,00010,045#19,00015,040#05,00018,041#05,00021
05#052,00023,129#05,00026,05#043,00029,31#086,00032,31#137,00036
045#31,00040,101#06,00045,
H1092479
H202
H31170
H4001
H50830,00000,
001#19,00004,19#030,00015,045#19,00021,040#05,00024,041#05,00027
05#052,00033,129#05,00038,05#043,00041,31#086,00046,31#137,00050
045#31,00053,101#06,00059,
```

Figure 2. Raw data from focal animal observation sessions. Three tests are separated by header information.

```
H1092479
H202
H31165
H416
H50800
4,4,4,4,4,4,7,7,6,6,1,2,4,7,7,7,
*
4,4,4,4,4,4,7,7,5,5,1,4,4,7,7,7,
*
4,7,7,5,6,6,4,4,4,7,7,4,4,1,2,4,
*
4,4,4,4,7,7,7,7,7,7,4,4,5,5,5,4,
*
4,4,4,4,4,4,4,4,7,7,7,5,5,5,4,4,
*
H1092479
H202
H31170
H416
H51000
7,7,7,7,4,4,4,7,4,4,4,4,1,2,4,4,
*
3,1,5,5,4,4,5,5,3,3,4,4,4,4,7,7,
*
4,4,4,4,4,4,3,4,4,7,7,4,7,7,7,7,
*
3,3,4,4,4,4,7,7,4,7,7,5,5,1,2,4,
*
4,4,4,4,7,7,4,4,5,5,1,4,5,2,1,4,
*
H1092479
H202
H31175
H416
H51400
4,4,4,4,7,7,7,5,5,1,2,4,4,7,7,7,
*
4,4,4,4,4,4,7,7,5,1,2,7,7,7,7,7,
*
4,4,4,4,4,4,4,7,7,7,5,1,2,6,6,4,
*
5,5,7,4,4,5,7,7,6,6,4,4,4,4,4,4,
*
4,4,4,5,1,2,6,6,7,7,7,4,7,7,7,7,
*
```

Figure 3. Raw data from group scan observation sessions. Three tests are separated by header information.

COMPUTER ACCESS

One of the strengths of the Datamyte system is the ease of transmitting data to a host computer (Torgerson, 1977). We interface the Datamyte directly to a Digital Equipment Corporation PDP-11/03 via a Lear Seigler ADM-3A CRT, using commercially available cables and connectors. The Datamyte is extremely flexible and can

be easily adapted to various types of computer terminals with no special hardware. This is in sharp contrast to other systems, which require extensive support hardware.

Data are transmitted from the Datamyte to the PDP-11/03 at 1,200 baud in ASCII code and are displayed simultaneously on the CRT. A FORTRAN program, DUMPER, controls file access on the PDP-11/03 and buffers the Datamyte output. This handshaking interface is necessary because the Datamyte, if uncontrolled, rapidly overflows the serial I/O buffers on the PDP-11/03, and data are lost. Upon completion of the data transfer, two stages of data editing begin.

DATA EDITING PROCEDURES

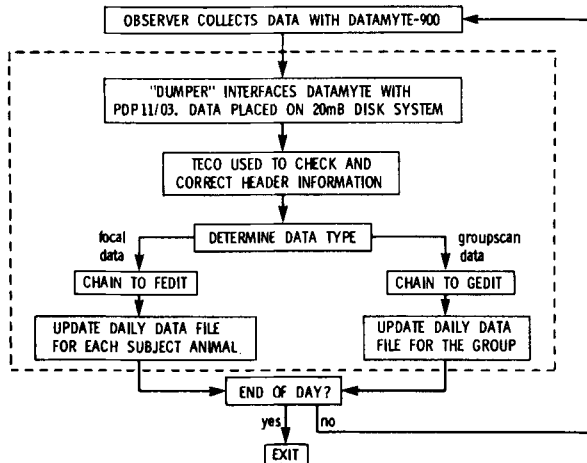
Two interactive programs are used to edit raw data and to put the data into a standardized format suitable for further analysis. The two programs are FEDIT, for focal observational data, and GEDIT, for group scan data.

After the transmission of data is complete, and before control is passed to the specific editing programs, the user can make specific changes in the data using the system text editor, TECO. Although under indirect command file control, the user has available all TECO editing capabilities and is expected to check the header information (see Figure 1) for both syntactical and logical correctness. Upon completion of editing using TECO, control is automatically passed to the appropriate editing/reformatting program. Figure 4 shows a general system flowchart.

FEDIT

This program first reads in a set of header informa-

BORES FLOWCHART



Note. Transfers within the dashed box are under indirect command file control.

Figure 4. Flowchart for data processing procedure.

FOCAL EDITOR FLOWCHART

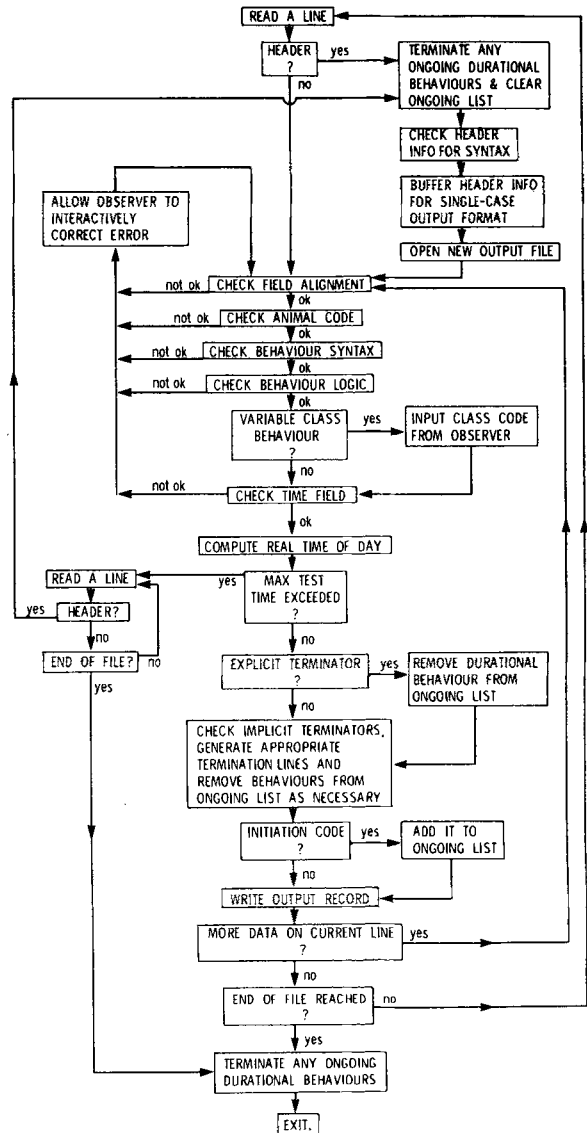


Figure 5. Flowchart for focal data editing procedure.

tion and, after checking its syntax, branches to a set of data editing modules. The edited and reformatted data are output to a file; the file's name is a function of the header information. Hence, we can tell by a file's name the subject of a test and the date of the test contained therein. Upon encountering a subsequent set of header information, the current output file is closed, a new one is opened, and processing continues. Figure 5 shows a flowchart of the focal data editor.

The first subroutine, ANIMCK, checks animal code numbers to be sure they are within the valid range of codes for the group and inserts the current interactant animal into a special buffer for future implicit terminator tests. If an error condition is detected, the location of the error and its type are printed on the CRT,

and the observer may then retype the entire data entry, insert appropriate characters, or delete the data entry. We are now collecting data on 37 animals, but this subroutine accommodates up to 99 individuals. Next, behavioral codes are checked for syntax and validity. The subroutine BHVSYN scans the data entry for a behavior and determines if it is within the valid range⁴ and correctly aligned in the data field. Misalignment occurs mainly from missed keystrokes during data entry. If an error condition is detected, it is handled as before.

The data record is then checked for logical inconsistencies. **BLOGIC** subroutine checks to see if the behavior is logically impossible or out of sequence. **BLOGIC** then determines the appropriateness of the interactant. For example, focal animal and recipient of the interaction must be identical if the behavior is self-directed (e.g., self-groom). This subroutine also assigns each behavior to a predetermined behavior class. The seven behavior classes are the same as the ones used in the group scan sessions. Of course, some behaviors may occur in multiple classes; for example, a hit can be Class 1 (aggression) or Class 5 (play), depending on the context. For behaviors that can be assigned to variable classes, **BLOGIC** calls another subroutine, **VARCLS**, which requires the observer to specify the class of the behavioral interaction. **VARCLS** checks the behavior class entered by the observer against predefined valid class codes for each variable-class behavior. When a legal behavior class has been assigned, **BLOGIC** is once again called to double-check the logic of the observer's choice. If logical inconsistencies are determined, the error is handled as before.

The time field of the observation is scanned next. Subroutine **TIMECK** checks the syntax of the field and then calculates the true time of day for an interaction by adding the local time field to the test's base start time given in the header information. If errors are detected, they are handled as before. Data are then checked by subroutine **EXPTRM** to see if the behavior encoded is an explicit terminator of a durational behavior. If it is, then the subroutine checks to see if there is, in fact, a valid ongoing durational behavior. If not, an error condition is detected, and the behavior is deleted. If it is a valid explicit terminator of an ongoing durational behavior, the behavior is removed from the list of currently ongoing durational behaviors. This program is designed to handle up to 10 instances of simultaneously occurring interactions for each durational behavior. For example, multiple types of durational behaviors can occur with several individuals simultaneously.

Next, the data entry is checked to see if it could be an implicit terminator of any ongoing durational behavior. Because there are occasions when an explicit terminator of a durational behavior may not be entered, subroutine **TERMCK** checks to see if the particular behavior logically terminates an ongoing durational behavior. This is an important feature and is useful in producing a com-

plete, logically consistent data set. The subroutine first checks to see if there is an ongoing durational behavior. If there is, the subroutine determines whether the focal animal is initiating or receiving that behavior. Next, it determines whether the interactants are the same in the ongoing and present behaviors, or if the behavior occurring involves another member of the group. If any condition exists that would logically terminate the ongoing durational behavior, an explicit termination code is generated with the associated time of the implicit terminator, and that interaction is removed from the list of ongoing durational behaviors. If conditions do not exist for logical termination, the durational behavior continues.

Finally, the behavior is checked to determine whether this interaction is the beginning of a new durational behavior. If so, it is added to the list of ongoing activities for future termination tests. After the data are checked, they are written to a file in a consistent reformatted form. Figure 6 provides an example of the reformatted focal animal observational data.

GEDIT

The **GEDIT** program is considerably simpler in editing procedures than **FEDIT**. After the data have been transferred to disk, the observer again has the opportunity to check and edit any of the header information that has been entered using the system editor, **TECO**. After the header information is edited, the program scans the data records and counts the number of animals that have been classified into each of the seven categories for a given group scan. If there are too many animals counted in a given session, it deletes those in excess of the total number of animals in the group. We assume that scoring errors are randomly distributed across behavioral classes and, as such, nullify each other over time. If too few are scored, those missed are classified as unknown. The program is designed for variable group sizes as entered in the header information. Figure 7 shows an example of the reformatted group scan data.

DATA SUMMARY

Focal Animal Data

The first nine columns of the focal data (Figure 6) are in the form of "who does what to whom?" For the first data record, Animal 01 initiates a 001 to Animal 19 (the recipient). Similarly, for the second record, Animal 19 initiates a 030 to Animal 01 (the recipient). Therefore, the initiator is in Columns 1-2, the behavior in Columns 4-6, and the recipient in Columns 8-9. Columns 11-16 contain the time of the interaction on a 24-h clock. Columns 18-23 contain the date of the observations (month, day, and year). The observer number is encoded in Columns 24-25. Columns 26-29 record the environmental data for that observation

Great consideration was given to adhering to the ANSI standard and to making the modules as portable as possible. Nonstandard features are modularized out of the body of the code and are marked as nonstandard. The modules are thoroughly documented, to facilitate modifications and to enhance understanding of the package as a whole. Although the algorithms used in these packages are applicable to most behavioral data sets, many of the individual aspects are designed around our particular research protocol and require modification before being applied to other particular data.

SUMMARY

The behavioral data collection and editing system (BORES) outlined in this paper makes extensive use of the capabilities of the Datamyte 900. We have found this unit to be highly reliable and extremely flexible. Unlike other data collection and analysis systems (Miller, Maginnis, & Wilson, 1979), our system offers considerable flexibility in the behaviors scored, the option of scoring multiple sessions of both types of data during a single day, and the option of increasing the number of animals sampled. The editing system checks the data for logical inconsistencies and syntax errors. Durations of behaviors are accurately handled through the use of implicit and explicit termination codes. Data are stored in a form that is suitable for a variety of other types of analyses.

REFERENCE NOTE

1. Peffer-Smith, P. G., & Smith, E. O. *Adult male-immature interactions in a captive stumptail macaque (Macaca arctoides) group*. Manuscript submitted for publication, 1979.

REFERENCES

ALTMANN, J. Observational study of behavior: Sampling methods. *Behaviour*, 1974, **49**, 227-265.
 CELHOFFER, L., BOUKYDIS, C., MINDE, K., & MUIR, E. The DCR-II event recorder: A portable high-speed digital cassette system with direct computer access. *Behavior Research Methods & Instrumentation*, 1977, **9**, 442-446.

COHEN, J. A coefficient of agreement for nominal scales. *Educational and Psychological Measures*, 1960, **20**, 37-46.
 DIXON, W. J. (Ed.). *BMD: Biomedical computer programs*. Los Angeles: University of California Press, 1977.
 DIXON, W. J., & BROWN, M. B. (Eds.), *BMD-79: Biomedical computer programs, P-series*. Los Angeles: University of California Press, 1979.
 FITZPATRICK, L. J. BEHAVE—An automated data analysis system for observed events. *Behavior Research Methods & Instrumentation*, 1977, **9**, 452-455.
 HOLLENBECK, A. R. Problems of reliability in behavioral research. In G. P. Sackett (Ed.), *Observing behavior: Data collection and analysis methods* (Vol. 2). Baltimore, Md: University Park Press, 1978.
 HOLM, R. A. Techniques of recording observational data. In G. P. Sackett (Ed.), *Observing behavior: Data collection and analysis methods* (Vol. 2). Baltimore, Md: University Park Press, 1978.
 MILLER, D. F., MAGINNIS, P. T., & WILSON, M. C. SOBEX: A timesharing data system for sociopharmacological experiments. *Behavior Research Methods & Instrumentation*, 1979, **11**, 461-463.
 NIE, N. H., HULL, C. H., JENKINS, J. G., STEINBRENNER, K., & BENT, D. H. *SPSS: Statistical package for the social sciences* (2nd ed.). New York: McGraw-Hill, 1975.
 SIDOWSKI, J. B. (Ed.). *Behavior Research Methods & Instrumentation* (Vol. 9). Austin, Tex: Psychonomic Society, 1977.
 STEPHENSON, G. R., & ROBERTS, T. W. The SSR system 7: A general encoding system with computerized transcription. *Behavior Research Methods & Instrumentation*, 1977, **9**, 434-441.
 TORGERSON, L. Datamyte 900. *Behavior Research Methods & Instrumentation*, 1977, **9**, 405-406.

NOTES

1. DEC, PDP, and TECO are registered trademarks of the Digital Equipment Corporation.
2. A 15-min observation session is of adequate duration.
3. Weather information is encoded in four digits. The first digit indicates cloud cover: 1 = clear, 2 = 25% overcast, 3 = 50% overcast, 4 = overcast, and 5 = rain. The second digit indicates wind conditions: 1 = calm, 2 = light, and 3 = gusty. The third and fourth digits indicate temperature in degrees Centigrade.
4. We are presently scoring 140 different behaviors. See Peffer-Smith and Smith (Note 1) for the details of the behavioral inventory. The editing program will handle up to 999 different behaviors.

(Received for publication January 8, 1980;
 revision accepted January 22, 1980.)