Comparing observational software with paper and pencil for time-sampled data: A field test of Interval Manager (INTMAN)

JON TAPP

Vanderbilt University, Nashville, Tennessee

and

RENATA TICHA, ERIN KRYZER, MEAGHAN GUSTAFSON,
MEGAN R. GUNNAR, and FRANK J. SYMONS

University of Minnesota, Minneapolis, Minnesota

In this article, we describe the Interval Manager (INTMAN) software system for collecting time-sampled observational data and present a preliminary application comparing the program with a traditional paper-and-pencil method. INTMAN is a computer-assisted alternative to traditional paper-and-pencil methods for collecting fixed interval time-sampled observational data. The INTMAN data collection software runs on Pocket PC handheld computers and includes a desktop application for Microsoft Windows that is used for data analysis. Standard analysis options include modified frequencies, percent of intervals, conditional probabilities, and kappa agreement matrices and values. INTMAN and a standardized paper-and-pencil method were compared under identical conditions on five dimensions: setup time, duration of data entry, duration of interobserver agreement calculations, accuracy, and cost. Overall, the computer-assisted program was a more efficient and accurate data collection system for time-sampled data than the traditional method.

INTMAN is an observational software system used for time-sampled data in behavioral research projects that rely on direct observation. Traditional methods of collecting this type of data have normally relied on using paper-and-pencil methods. In the past 15 years, computer-assisted direct observation programs have become widely available and readily accessible with increasingly affordable hardware options (Thompson, Felce, & Symons, 2000) applied in a wide range of educational settings (Greenwood, Carta, & Dawson, 2000) and community-based settings (Felce & Emerson, 2000). Despite claims suggesting or implying that computer-assisted data collection is superior to traditional methods, to date there is a very limited number of studies that empirically compare the two methods of data collection in general or specific to time-sampling procedures on dimensions relevant to research considerations including efficiency, accuracy, and cost.

White, King, and Duncan (2002) compared a traditional data collection method in the form of paper and

pencil with voice recognition technology (VRT) in a study of social interaction and singing in cowbirds. The time for data input and interobserver agreement were contrasted, and overall results showed that the VRT method was more time efficient and led to significantly more accurate interobserver agreement. Saudargas and Zanolli (1990) and Klesges, Woolfrey, and Vollmer (1985) contrasted traditional and computer-assisted methods to assess the accuracy of momentary time sampling against real-time data collection but did not focus on the advantages and disadvantages of the methods themselves along a broader set of dimensions relevant to the decisions a researcher would need to make in choosing between traditional and computer-assisted methods.

The purpose of this article is twofold. The first is to introduce and describe the Interval Manager (INTMAN) software system designed to facilitate time-sampled data collection. The second is to report the results of a preliminary study that directly compared INTMAN with a traditional paper-and-pencil time-sampled data collection system along five dimensions including setup time, duration of data entry, duration of interobserver agreement calculations based on Cohen's kappa, accuracy, and cost.

INTMAN: System Description and Program Operation

The INTMAN software system was designed to facilitate the collection of time-sampled data in field-based settings. INTMAN has a collection component and an analysis component. The collection component runs on Pocket

The first author and the first author's not-for-profit institution receive funds from license sales and therefore have a financial interest in the software described in this article. The research was supported in part by NICHD Grant 36582 and NIMH Grant 62601 to the University of Minnesota, as well as a McKnight Land-Grant Professorship to F. J. Symons. The authors thank John Hoch for his technical assistance through the Observational Methods Lab at the University of Minnesota. Correspondence should be addressed to J. Tapp, Vanderbilt University, Vanderbilt Kennedy Center, Peabody Box 74, 230 Appleton Place, Nashville, TN 37203 (e-mail: jon.tapp@vanderbilt.edu).

PC devices that use Windows CE. The analysis component runs on computers with the Windows 98 or higher operating system. The minimum hardware requirements for the operating system are adequate for INTMAN. Data are collected with the handheld computer and transferred to the desktop computer for storage and analysis. The programs are operated using standard Windows controls and dialog boxes.

The INTMAN software uses the clock in the handheld computer and a series of drop-down boxes on the touch screen of the handheld computer. The time interval can be selected and is given in two values: an "observe" period and a "record" period. The state of the codes on the screen is recorded at the end of the record period. This yields a data file with a column of codes for each category for each interval of time in the session. Session duration is determined by preselecting the number of time intervals to collect. Codes are grouped into categories, and each category needs to be mutually exclusive, since two codes cannot be selected simultaneously. INTMAN can display and collect up to 22 categories of codes with up to 20 codes within each category.

To control what is displayed on the handheld screen, a code file is used. Code files in INTMAN are made with a standard text editor, such as Windows Notepad. The code file contains the codes, data numbers, and screen labels for the codes. As an example, consider a simple direct observational data collection scenario with a simple three-category coding system designed to support studying aggression in schools. One specific research question may address the issue of whether aggression varies as a function of setting and the teacher's presence or absence. Possible codes for the first category would likely include (1) aggression occurred (during the "observe" period) and (2) no aggression occurred (during the "observe" period). The second category for settings may include (1) classroom, (2) playground, and (3) transition. The third category would then be (1) teacher present and (2) teacher absent. Each of these categories is displayed on the screen as a selection in a drop-down list box for each category. When the clock is started, the computer plays the "observe" sound. The observer watches the class until the computer plays the "record" sound. The observer would select which state each category was in during the observe period; then, at the end of the record period, the states of the drop-down boxes are recorded into the data file along with the time.

Analysis Options

After the data are collected, they are transferred to the desktop PC for analysis. Three analyses are available: (1) frequency analysis, (2) interobserver agreement, and (3) conditional probabilities. The analysis program works using the concept of a "data set." The data set consists of results files (output) with an attached set of observation data files. The analysis software processes the observation data into a result file. The data set is created by using the toolbar in the analysis program to name and select the files.

The frequency analysis processes the list of observation files into a result file that consists of total frequencies of

all the codes for each category. The sums can be displayed as totals for each file and/or as a pooled sum for all the files in the result set. The number of times each code was scored and the proportion of the total that number represents are displayed in the result file. Result files can be formatted either as text output or as comma-separated values for using in statistical and spreadsheet programs.

The interobserver agreement analysis tallies agreements interval by interval for each category. A kappa matrix (see Cohen, 1960) is provided to allow easy display of disagreements during a session. Cohen's kappa and an agreement ratio ($A/(A+D)$) are calculated as indices of agreement, and sessions can be pooled providing overall agreement numbers.

Conditional probabilities count code frequencies of pairs of categories considered together. Using the above coding scheme as an example, this analysis would provide the number of intervals in which aggression occurred with teacher present compared with the number of intervals in which aggression occurred with teacher absent. The observed probability of aggression as a function of teacher presence would be supplied in the results using the frequency counts.

Related Features

The INTMAN program can also analyze data collected from digital video using the ProCoderDV system (PCDV; Tapp, 2003) and data collected from an older DOS version of INTMAN (Tapp, Youngquist, & Odom, 1993). Data sets of files can be saved for later use. The analysis program also has a built-in data file editor that allows for editing of the data using the screen labels rather than their numeric representation from the data file.

APPLICATION

Method

To field-test INTMAN, we compared it directly with the Modified Observational Ratings of Caregiver Environment (M-ORCE), a standardized and validated paper-and-pencil direct observational protocol for interval coding (see below; see Appendix A for the paper version of the coding form and Appendix B for the INTMAN screen of the coding form). To conduct the comparison, two trained data collectors coded three 10-min video segments from standard video-collected data of children in home-based child-care environments. Observers coded each video segment with M-ORCE then with INTMAN using Pocket PCs, or vice versa (order was randomized). The two methods were compared on five dimensions including setup time, duration of data entry, duration of kappa calculations, accuracy, and cost (see White et al., 2002).

The M-ORCE was developed for rating quantitative and qualitative child-care variables relevant to early child care (Gunnar & Kryzer, 2001). The M-ORCE is an instrument that examines the characteristics of the child's experiences in the child-care setting by recording the frequency of specific kinds of behavior directed by the caregiver and other children toward the target child and specific behaviors displayed by the target child, by rating the quality of the caregiver's behavior and the target child's behavior in the environment, and by rating the overall climate of the child-care setting. The behavior (frequency) scales record the occurrence and/or level of key behaviors in the following areas: language-focused interaction, stimulation of development, behavior management, child's activity, child's interaction(s) with other children, and child's compliance. The M-ORCE takes 44 min to complete, 30 min of which are

used for the behavior scales. The 30 min are broken down into three 10-min segments, during which coders watch the target child's behavior for 30 sec and then make notes of the child's behavior for 30 sec. This formed the basis of selecting three 10-min segments for comparison purposes.

Results

The results reported are comparisons per subject. In terms of session setup time, following initial setup for either system, both systems were brief (~20 sec). It should be noted, however, that considerable time was spent in setting up the initial working version of the M-ORCE in terms of code arrangement within boxes for observers to score and so on, whereas this process was automated through INTMAN by virtue of a screen array with pull-down windows corresponding to codes. In terms of duration of data entry, it took significantly less time to enter observational data into SPSS using INTMAN in comparison with the paper-and-pencil method (1 min 30 sec vs. 30 min, respectively). The time devoted to calculating Cohen's kappa coefficients for interobserver agreement checks was considerably less using INTMAN (40 sec) than with paper and pencil (45 min). We also compared INTMAN kappa calculation time with a standard program for calculating kappa (ComKappa; Robinson & Bakeman, 1998) in which raw data were entered into the program (15 min). There was no degradation in the accuracy of the data collected via INTMAN relative to paper and pencil. Accuracy was defined as the precision of interobserver agreement calculated using kappa from the three 10-min videotape segments. Overall kappa was .90 for INTMAN and .82 for M-ORCE. Finally, related to cost, there was an initial investment when using INTMAN (\$999) and the Pocket PC with accessories (\$400) that was significantly more than the investment into paper-and-pencil supplies. On a longer time scale, however, an average of 1 h 13 min of work was saved per subject per observation session. On the basis of an average wage of \$14 per hour, approximately \$17 would be saved in salary per observation session per subject. A project with 20 subjects with 20 observation points would save approximately \$6,800.

DISCUSSION

In this article, INTMAN, an observational software data-collection program for time-sampled data, was introduced and described. Key features included flexibility in determining the array and number of codes available, as well as analyses options that included modified frequencies, percent of intervals, conditional probabilities, and Cohen's kappa coefficients to index observer agreement. There were minor costs associated with start-up (hardware purchase) and an upper limit in the number of categories of codes available (22; constrained by screen size and pull-down menu bars), but within each category, 20 unique codes could be programmed.

Because few direct comparisons of computer-assisted and traditional paper-and-pencil methods exist, we conducted a preliminary test and contrasted INTMAN with an estab-

lished standardized paper-and-pencil time-sampled system (M-ORCE). The computer-assisted method INTMAN was more time efficient, more accurate, and more cost efficient over a long time scale than the traditional paper-and-pencil version for time-sampled data collection. On this last point, when surveying the options available among observation systems and collection methods, it may be worthwhile to consider the timescale of a given project and any future planned projects. It would seem reasonable (1) that if any one project may include repeated measures over some appreciable timescale (e.g., weeks) and (2) that if any one project's study sample exceeds 5 or more subjects, then (3) some consideration should be given to an initial investment in observational software and hardware as a cost-effective and efficient alternative to traditional paper-and-pencil methods (Smith, Tapp, & Warren, 2000).

In summary, technological innovations can provide much needed assistance for researchers using direct observation in behavioral research designed to quantify behavior in real-world contexts. In this preliminary application, it appeared that computer-assisted methodology outperformed traditional paper-and-pencil methodology along several dimensions important to broader research considerations. As with all applications of technology, however, it may be worth stating here that utility is just as often determined by clear guiding research questions with correspondingly clear analytic plans. Perhaps more so than with paper and pencil, computer-assisted technologies make it possible to produce an abundance of data in the absence of either. Technology can support and enhance, but not replace, research design.

REFERENCES

- COHEN, J. (1960). A coefficient of agreement for nominal scales. *Educational & Psychological Measurement*, **20**, 37-46.
- FELCE, D., & EMERSON, E. (2000). Observational methods in assessment of quality of life. In T. Thompson, D. Felce, & F. J. Symons (Eds.), *Behavioral observation: Technology and applications in developmental disabilities* (pp. 159-176). Baltimore: Paul H. Brookes.
- GREENWOOD, C. R., CARTA, J. J., & DAWSON, H. (2000). Ecobehavioral assessment systems software (EBASS): A system for observation in education settings. In T. Thompson, D. Felce, & F. J. Symons (Eds.), *Behavioral observation: Technology and applications in developmental disabilities* (pp. 229-252). Baltimore: Paul H. Brookes.
- GUNNAR, M. R., & KRYZER, E. (2001). *Modified Observational Ratings of Caregiving Environment (M-ORCE)*. University of Minnesota: Institute of Child Development.
- KLESGES, R. C., WOOLFREY, J., & VOLLMER, J. (1985). An evaluation of the reliability of time sampling versus continuous observation data collection. *Journal of Behavior Therapy & Experimental Psychiatry*, **16**, 303-307.
- ROBINSON, B. F., & BAKEMAN, R. (1998). ComKappa: A Windows 95 program for calculating kappa and related statistics. *Behavior Research Methods, Instruments, & Computers*, **30**, 731-732.
- SAUDARGAS, R. A., & ZANOLLI, K. (1990). Momentary time sampling as an estimate of percentage time: A field validation. *Journal of Applied Analysis*, **23**, 533-537.
- SMITH, J., TAPP, J., & WARREN, S. F. (2000). Analysis of early communication and language intervention practices using observational technology. In T. Thompson, D. Felce, & F. J. Symons (Eds.), *Behavioral observation: Technology and applications in developmental disabilities* (pp. 215-227). Baltimore: Paul H. Brookes.

TAPP, J. (2003). ProCoder for Digital Video [Computer software]. Nashville, TN: Vanderbilt Kennedy Center.

TAPP, J., YOUNGQUIST, G., & ODOM, S. (1993). INTMAN for DOS [Computer software]. Nashville, TN: Vanderbilt Kennedy Center.

THOMPSON, T., FELCE, D., & SYMONS, F. J. (Eds.) (2000). *Behavioral ob-*

servation: Technology and applications in developmental disabilities. Baltimore: Paul H. Brookes.

WHITE, D. J., KING, A. P., & DUNCAN, S. D. (2002). Voice recognition technology as a tool for behavioral research. *Behavior Research Methods, Instruments, & Computers*, **34**, 1-5.

APPENDIX A

M-ORCE Coding Form

Subject _____	Subject Number _____	Obs _____	Date _____	Time _____						
ADULT (S) AVAILABLE	Beginning of cycle	End of cycle	CHILDREN AVAILABLE	Beginning of cycle	End of cycle					
Caregiver 1	_____	_____	Number of infants/toddlers (# of)	_____	_____					
Caregiver 2	_____	_____	Preschoolers (# of)	_____	_____					
Caregiver 3	_____	_____	Grade schoolers (# of)	_____	_____					
Other adult(s) (# of)	_____	_____								
INTERVAL	21	22	23	24	25	26	27	28	29	30
CHILD'S ACTIVITY CONTEXT	21	22	23	24	25	26	27	28	29	30
Physical Care										
Purposeful Transition/Waiting										
Watching TV										
Time Out										
Watching/Wandering/Unoccupied										
Solitary play/activity										
Organized group activity										
Other activity										
CHILD'S ATTENTION/ENGAGEMENT-AVERAGE	21	22	23	24	25	26	27	28	29	30
Uninvolved (U); Low(L); High(H)										
ADULT LANGUAGE	21	22	23	24	25	26	27	28	29	30
Speaks positively/neutral to child/ren										
Speaks negatively to child/ren										
ADULT STIMULATION	21	22	23	24	25	26	27	28	29	30
Teaches academic rule										
Teaches social rule										
Teaches other skills/information										
Reads/tells story/sings/plays game										
Supervises project (arts/crafts/other)										
Positive physical contact										
Mutual exchange										
ADULT PHYSICAL CONTROL	21	22	23	24	25	26	27	28	29	30
Restrictive actions										
C'S LEVEL OF POS/NEU SOCIAL INTEGRATION-PREDOMINANT	21	22	23	24	25	26	27	28	29	30
Unintegrated (U); Low(L); High(H)										
CHILD'S NEGATIVE INTERACTIONS	21	22	23	24	25	26	27	28	29	30
Not Applicable										
Rejected/rebuffed/or ignored in attempt										
Directs negative social action										
Receives negative social action										
Adult intervention										
C'S COMPLIANCE WITH ADULT DIRECTION	21	22	23	24	25	26	27	28	29	30
Not Applicable										
"CR", "CS", "I", "A", "R"										

APPENDIX B
INTMAN Coding Screen

Interval Manager

Data File not Open...

"Physic"	N	"Reads"	N
"Transi"	N	"Superv"	N
"TV"	N	"Posit"	N
"Time"	N	"Mutual"	N
"Watch"	N	"Restrict"	N
"Activ"	N	"Integr"	
"OtherA"	Solitary	"Reject"	N
"Child's"	Group	"Directs"	N
"A.L."	S+G	"Receiv"	N
"Teach"	N	"Adult"	N
"OtherT"	N	"Compl"	U

Pick new to start a new file. Interval:

File Tools

(Manuscript received August 16, 2004;
revision accepted for publication February 7, 2005.)