Behavioral Interventions
Behav. Intervent. **19**: 275–285 (2004)
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI**: 10.1002/bin.159

A REVIEW OF THE BEHAVIORAL EVALUATION STRATEGY AND TAXONOMY (BEST®) SOFTWARE APPLICATION

Tina M. Sidener, Daniel B. Shabani and James E. Carr*
Western Michigan University, Kalamazoo, MI, USA

Recent computer technology has led to the development of a number of software applications that have been specifically designed for collecting and analyzing observational data in real time. Behavioral Evaluation Strategy and Taxonomy (BEST®) is an innovative software program that provides users with an effective way to collect, store, and analyze real-time observational data. The program is comprised of two distinct applications: BEST Collection® and BEST Analysis®. The purpose of the current article was to provide a critical review of BEST Version 4.1.6 for the Windows® (95/98/NT) operating system. The basis of this review was our use of BEST to collect and analyze data for several studies over a 2 year period. Copyright © 2004 John Wiley & Sons, Ltd.

INTRODUCTION

Applied behavior analysis has long been associated with the rigorous and repeated measurement of behavior (Baer, Wolf, & Risley, 1968). Direct-observation data collection procedures such as interval recording and momentary time sampling arose from this tradition. Consequently, a number of data-analysis techniques (e.g. exact agreement, Cohen's kappa) have been developed to assess the reliability of the products of data collection. In addition, sophisticated techniques have been developed to analyze local behavior patterns within session. The combination of relatively demanding data collection and analysis procedures can sometimes prove burdensome, especially when data are collected using paper-and-pencil scoring methods and subsequently analyzed by hand. Fortunately, advances in computer technology have led to the development of a number of software programs that have been designed specifically for collecting and analyzing observational data in real time, eliminating some of the response effort associated with manual alternatives. Kahng and Iwata (1998, 2000) recently published much-needed reviews that

*Correspondence to: James E. Carr, Department of Psychology, Western Michigan University, 1903 W. Michigan Avenue, Kalamazoo, MI 49008-5439, USA. E-mail: jim.carr@wmich.edu

presented the general characteristics of a number of these software applications, including Behavioral Evaluation Strategy and Taxonomy (BEST®). BEST is described in its accompanying technical manual as 'a user-friendly software package that facilitates the real-time collection and analysis of observational category system data' (Sharpe & Koperwas, 1999, back cover).

The purpose of the current article is to extend the information reported by Kahng and Iwata (1998, 2000) by providing a more comprehensive and critical review of the BEST application. The basis of the review was our use of BEST to collect and analyze data for several studies over a 2 year period. During the review process, we were neither creatively nor financially affiliated with the software developers.

We reviewed BEST version 4.1.6 for the Windows® (95/98/NT) operating system. The system requirements for BEST include at least a 386 MHz processor and 32 MB of memory. The current version of BEST is available from Educational Consulting, Inc. (http://www.skware.com). Individual software units cost $499.00 and each additional unit costs $100.00. Site licenses, which include five units and free upgrades, cost $999.00, with additional units costing $100.00. BEST can also be licensed to entire universities for $2000.00 per year.

BEST is comprised of two distinct applications: BEST Collection® and BEST Analysis®. Thus, our review is organized into two primary sections, followed by some general comments. Because BEST has dozens of capabilities, including sophisticated analysis functions, a comprehensive review of all of them is beyond the scope of this article. Therefore, our review is based primarily on conventional usage of the program.

DATA COLLECTION

The BEST Collection program allows the user to record real-time data via several input devices (e.g. keyboard, microswitch connected to the serial port). In order to begin collecting data, the first step is the customization of an observation template that will meet the needs of the specific data collection project.

Configuration Files

An alphanumeric Keyboard Configuration graphic with a likeness of a computer keyboard is used for programming *configuration files* (i.e. data collection templates). After a new configuration file has been opened, the user may assign up to 36 different behaviors to the number and letter keys of the keyboard. Clicking on a key opens the Key Configuration submenu (Figure 1), which allows for various programming options regarding that behavior. For example, the user may specify an acronym for

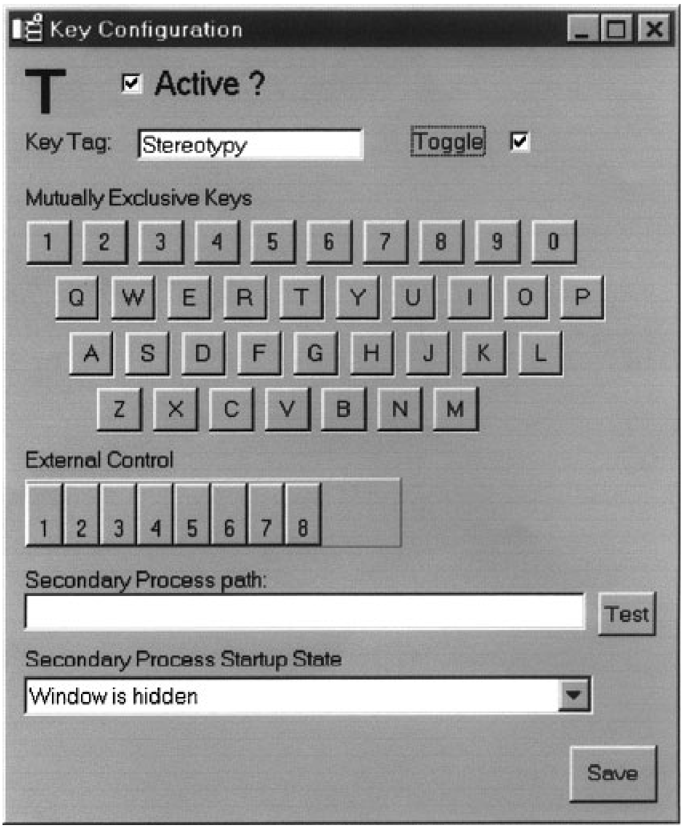

Figure 1. Key Configuration submenu that is used to program the keys within the configuration file.

the key, identify an external data collection device, or activate secondary software applications such as a word processor. Assigned key-tag names appear on the keys, allowing for easier identification during collection. The toggle option allows the user to specify frequency or duration recording for each key. There is currently an option to 'activate' or turn on each key after programming it (i.e. naming and specifying its usage). A more functional alternative, however, might be a box that could be deselected to render it inactive—or the option could be totally eliminated. Most users would likely only program keys they intend to use as active keys.

Timers may be programmed in seconds and provide features useful to most researchers. However, the terms used to describe their functions do not lend themselves to easy usage without careful and thorough readings of the descriptions in the manual. Perhaps more self-explanatory terms that describe the visual and automatic termination features would be helpful. For example, the essential feature of the *Count Down Timer* is terminating data collection. When the *Count Down Timer* expires, no more data may be collected. This distinguishes it from the *Seconds Timer*, which merely provides a running numerical count in seconds, and does not terminate data collection. Therefore more self-explanatory labels might be 'Countdown Termination Timer' and 'Count-Up Timer', respectively.

After configuration files are created, they are saved to the hard drive, and may be used repeatedly throughout a research or applied project. In addition, an unlimited number of configuration files may be created, allowing for the development of a variety of observation templates that may be used in a number of different situations.

Data Files

To begin collecting data, the user must first open a configuration file, and then open a new data file. Data are collected via the Event Recorder Screen (Figure 2), which is a user-friendly keyboard graphic similar to the one used to set up the configuration file. Pressing or clicking on a key programmed for a specific event highlights the key and records time of occurrence, frequency, and duration information about the event. Multiple keys may be pressed simultaneously, allowing for precise recording of behavioral events in relation to one another. The *Start, End, Pause*, and *Options*

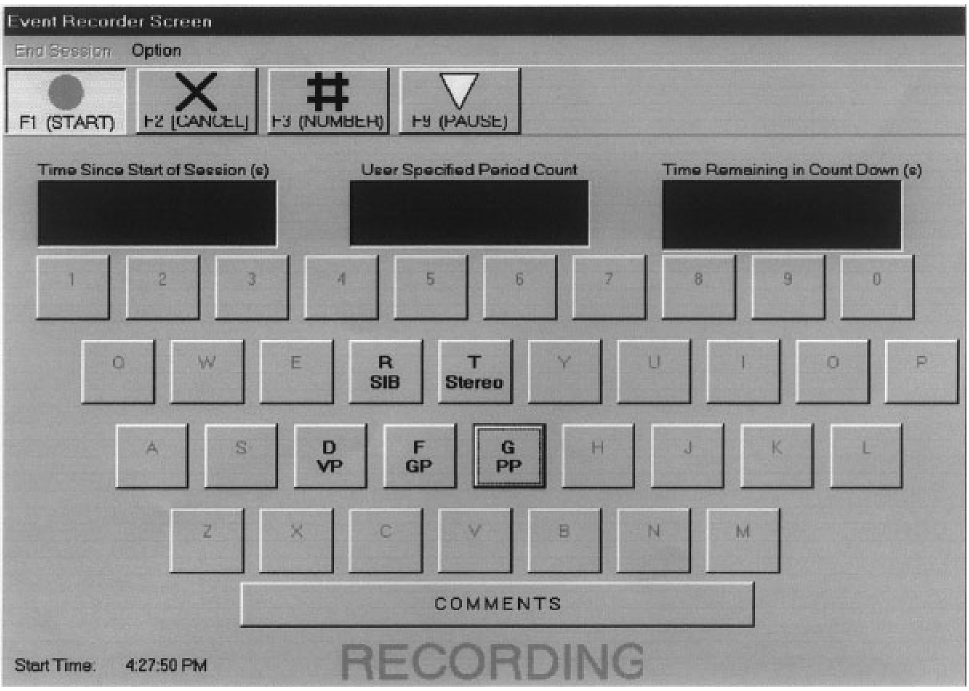

Figure 2. Event Recorder Screen. This keyboard graphic is visible on screen during data collection.

buttons are self-instructional, and could be easily used without reference to the manual. The *Cancel* option is a timesaving and practical feature, serving an 'erase' function. For example, if a key is inadvertently pressed, *Cancel* allows for the errant data to be deleted. In addition, qualitative information can be added to each data file, which is helpful when interesting or novel events occur but were not pre-identified (e.g. change in behavioral intensity, the occurrence of a relevant environmental event). This feature allows for a more complete description of the observational setting.

Future versions of BEST Collection might include more self-instructional prompts for opening, saving, and indicating data files to be used. If, for example, the user does not want to save the data that have been collected and wants to record over the data file, there is currently no option to perform this operation. Although a prompt appears asking whether the user wants to save the file, choosing 'no' does not allow the user to rerecord over the existing file. The only way to record over an old data file or enter in a new data file is to exit the data collection program and select a preexisting data file or create a new data collection file.

During data collection, two additional windows may be opened to allow the user to monitor the frequency and duration of behaviors. This feature may be useful when termination of an observation session is contingent upon the occurrence of a specific number or duration of events (e.g. self-injury occurring 25 times or for more than 30 s). Unfortunately, the additional windows cover part of the keyboard graphic and thereby make it difficult to see all of the keys. In addition, the size of the windows cannot be adjusted and instead must be moved around until the user can establish a layout that allows for the best possible view of the keyboard graphic and the additional windows. Finally, there is currently no way to close these windows if *End Session* has already been selected. A close box would allow the user to close these windows after the keyboard graphic has already been closed.

BEST ANALYSIS

After data have been entered into the computer using BEST Collection, the resulting data file can then be analyzed using BEST Analysis. The data files created within the configuration program must be specified and grouped in the defined portion of the analysis window (e.g. data collection file path or sub-grouping file path). Individual data files or folders containing multiple data files may be entered and processed simultaneously in the data analysis window (Figure 3). Analysis options include a qualitative summary, a hierarchical presentation of quantitative information (e.g. frequency, duration, latency), sequential analysis (e.g. Z-scores, conditional probabilities), and visual illustrations (e.g. tables, graphs).

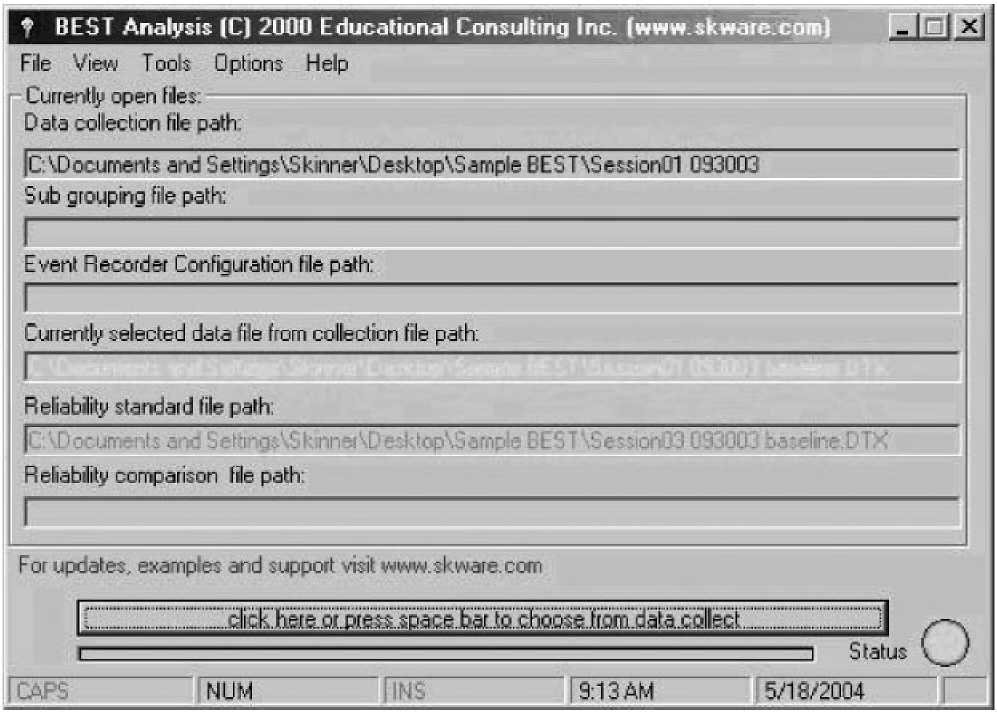

Figure 3. Data analysis window in which files are entered for data analysis and interobserver agreement assessment.

The results displayed in the Hierarchical View presentation (Figure 4) are conveniently separated into individual directories that include quantitative (e.g. descriptive information including frequency, duration, and mean) and qualitative (e.g. comments made during data collection) information about each particular behavior or event. Within-session graphs, multiple bar and pie chart comparisons, and grouping analyses are additional options available for analysis. A separate analysis program computes interobserver agreement (IOA) and includes total agreement, kappa, and overall point-by-point agreement. This is one of the most important advantages of using the BEST system; it provides an extremely efficient alternative to paper-and-pencil IOA calculations. IOA is calculated by entering a primary data file (i.e. reliability standard file) and a comparison file into the data analysis window (Figure 3). After the user has specified two data files for comparison, the user must specify several reliability parameters, including *Align Data Sets?* in order to synchronize the start times of the two data files being compared (Figure 5). An event

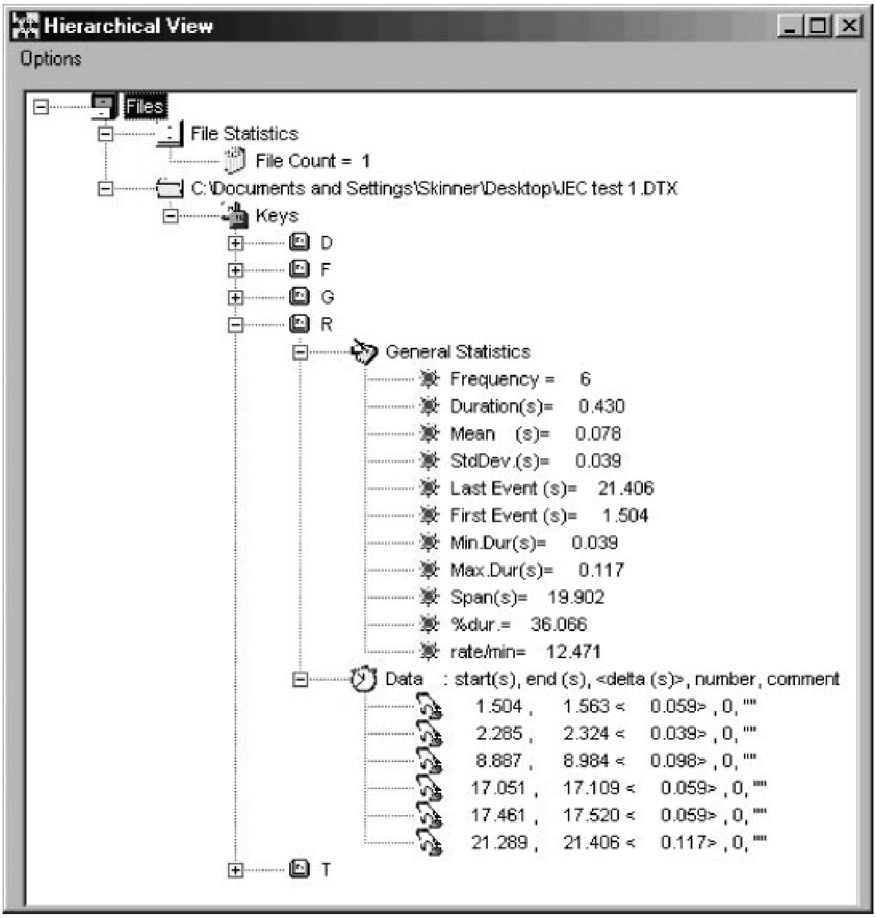

Figure 4. The Hierarchical View is the BEST Analysis program's default data summary.

tolerance option allows the user to program a time allowance between events, which serves as a range of acceptability for behaviors that occur at approximately the same time. For example, if 3 is entered for event tolerance, behaviors that occur within 3 s of each another will be counted as an agreement. This allows for the calculation of point-by-point agreement on the occurrence of each event without the need for partial- or whole-interval recording. The event duration shift option allows the user to specify whether single or multiple occurrences of the same behavior are recorded as an agreement or disagreement. For example, if one observer recorded the duration of toy play as 1 min and a second observer recorded toy play as three consecutive 20 s instances, event duration shift records these as one agreement instead of one agreement and two disagreements.

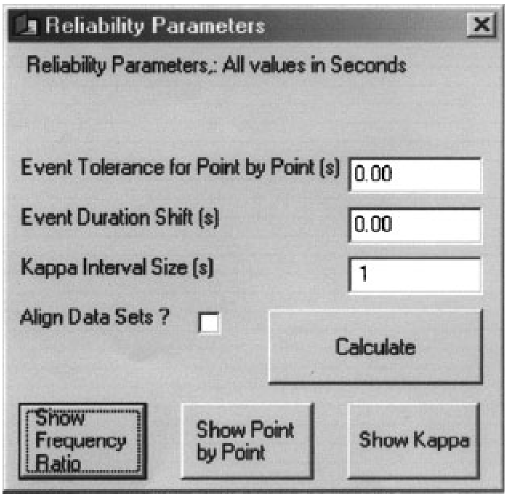

Figure 5. Reliability Parameters screen for interobserver agreement functions.

One difficulty we encountered when using software to record real-time observational data was the incongruence between real-time data collection and interval-based recording. Interval and time sampling recording procedures employ periods of time in which judgments are made about behavioral occurrence. During subsequent IOA assessment, these intervals can be compared on a point-by-point basis. However, when using a software application such as BEST, data are not collected within intervals; each event is recorded as it occurs. We experienced some initial confusion regarding IOA calculation (using BEST Analysis) using kappa and point-by-point agreement because data were not recorded in intervals. However, BEST allows the user to specify the interval to use for the calculation of IOA (from 1 s). So, although data are not collected in intervals, they are still analyzed within them.

Despite numerous benefits, there are some limitations of BEST Analysis associated with the manner in which data are displayed. First, although the configuration file provides the user with sufficient options to customize the observation template, the analysis program does not have this option. The information provided about a particular observation session is pre-determined and an option to customize which data are presented is not available. A customization option that would allow the user to specify the information he or she wants would be beneficial. For example, if the user is only interested in frequency and duration, then an option to provide only this information could be included in the set-up of the configuration file. Information regarding the standard deviation, percent duration, and the beginning and end time of each response may be unnecessary and distract from the primary variables of interest. Second, the manner in which the data are displayed is relatively elementary given the computer technology and graphics available for visual display of information. The data may be better analyzed through presentation in a table where the primary dependent measures (as specified by the user) are provided, instead of through a hierarchical presentation. Printing the hierarchical view of information has also proved cumbersome in that only the page displayed within the hierarchical view window (see Figure 4) can be printed. Hierarchical displays that spread across multiple pages require the user to view and print each page individually. Finally, important information in which a user would typically be interested is missing from the hierarchical presentation. Specifically, the names assigned to the target behaviors in the configuration file are not displayed in the hierarchical presentation of the data unless the configuration file with which the data were collected is also identified before conducting the analysis. This added step to conducting an analysis appears unnecessary given that it is difficult to think of a time when one would evaluate data from a particular observation template using a different data collection system than the one with which the actual data were recorded. Further, although the manual specified the aforementioned procedure for including behavior codes in the hierarchical display, we were unable to accomplish this using the configuration files we created. The only way we were able to label the target behaviors displayed in the hierarchical presentation was to customize the sample configuration files that were included with the BEST system (e.g. PeTeach). In addition, session length and interresponse times (IRT), which are commonly used, are not specified in the hierarchical display of data. Instead, session length is displayed under the statistics option for data analysis (i.e. statistics → tables → time span) and IRT must be calculated by hand using the individual start and end times provided in the hiera-rchical display. Both of these analysis options would be better served as part of a customized table that included specific measures in which the user was interested.

CONCLUDING REMARKS

Although we have not availed ourselves of all of the functions in BEST Analysis, the application offers an impressive array of additional possibilities, including sequential analysis (via the conditional probability index), within-session analysis (via graphical presentation), and fairly sophisticated sequential modeling functions. In addition, system requirements for BEST allow it to be used on most standard computers. The cost of BEST might be prohibitively expensive for individuals, but is within reason for universities and service agencies.

Available as a separate application is the BEST External® program, which allows peripheral devices to be connected to the computer's serial port through which data are subsequently recorded by BEST Collection. This application may be of particular interest to basic researchers, who could connect the BEST application to a variety of microswitches that could serve as manual operants (see, e.g., Lancioni, O'Reilly, & Basili, 2001). For example, a user could connect the levers in an operant chamber to the computer via the serial ports. BEST Collection could then be activated during sessions to automatically record the frequency and duration of, for example, lever presses, eliminating the need for a keyboard and a constant observer. In addition, another separate application, BEST Video®, allows videotape synchronization for video-based recording. There is also a version of BEST (BEST PPC) available for the PocketPC®.

We have generally found the BEST application to be a useful contribution to our research. The fact that it was written for Windows and, thus, has a graphic user interface is particularly appealing. BEST has been particularly helpful in simultaneously recording multiple topographies and for quickly calculating IOA. Overall, the program is an effective way to collect, store, and analyze observational data; however, several relatively minor revisions would render this software more user friendly. Learning how to effectively use BEST required repeated exposure to developing and using various data collection templates (i.e. configuration files) and repeated use of the data analysis program. In addition, we have experienced several bugs in the program (described above). However, these minor problems allowed us to come in contact with one of BEST's most impressive features: customer service. BEST was created by a behavior analyst and a software developer, both of whom have been directly available for technical assistance, most of which has been exemplary. Another nice feature of BEST is its technical manual. Although the manual provides many details, in some places the manual's descriptions could be clearer, and an index would be very helpful. Interestingly, the software also comes with a file-based Adobe Acrobat manual, although it is significantly different from the printed technical manual. The file begins with interesting discussions of Kantor's interbehavioral field theory and the implications of computers in data collection, as well as examples of behavior codes from disciplines such as special education and ethology. The file also includes information on staff training. Although we found the computerized manual interesting, we hope that its functions and relation to the printed manual are more clearly explicated for the user in future editions. We should note that because our version of BEST was a relatively early one, we expect improvements to software and technical materials to be a natural development across subsequent versions.

REFERENCES

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91–97.

Kahng, S., & Iwata, B. A. (1998). Computerized systems for collecting real-time observational data. *Journal of Applied Behavior Analysis, 31*, 253–261.

Kahng, S., & Iwata, B. A. (2000). Computer systems for collecting real-time observational data. In T. Thompson, D. Felce, & F. J. Symons (Eds.), *Behavioral observation: Technology and applications in developmental disabilities* (pp. 35–45). Baltimore, MD: Brookes.

Lancioni, G. E., O'Reilly, M. F., & Basili, G. (2001). Use of microswitches and speech output systems with people with severe/profound intellectual or multiple disabilities: A literature review. *Research in Developmental Disabilities, 22*, 21–40.

Sharpe, T., & Koperwas, J. (1999). *BEST®: Behavioral Evaluation Strategy and Taxonomy software* [Technical Manual]. Thousand Oaks, CA: Sage/Scolari.

