‘ECOVRD’ –A Tablet PC-Based Tool to Support Observational Studies

Ute Fischer, Deepak Jagdish, Abbas Attarwala
Georgia Institute of Technology

Abstract: This paper introduces a new Tablet PC-based tool (called ‘ECOVRD,’ Event Coding and Visual Representation of Data) to facilitate real-time classification of live or video-recorded individual and team behavior. The tool is easy to use as its design exploits the possibilities that the touch-based interface of the tablet PC provides. ECOVRD is highly versatile: coding categories are user-defined and can capture instances of behavior or the duration of behavior. In addition, a note-pad feature enables users to enter observations as text. ECOVRD also includes a visualization module that provides users with graphic data summaries and basic statistics for immediate feedback. Thus, it will be useful both to the research community and in applied settings, for instance to support professional coaching or training. This paper discusses the salient design and interaction aspects of ECOVRD, and its improvements over existing systems.

INTRODUCTION

Many studies in cognitive engineering and naturalistic decision making (NDM) involve observational data. To understand “how experienced people actually make decisions” (Zsambok, 1997, p. 4), researchers examine the behavior of professionals in their operational environments (e.g., Badke-Schaub, in press; Durso, Crutchfield, & Harvey, 2007; Schraagen & Leijenhorst, 2001; Lauche & Bayerl, in press; Patterson & Woods, 2001; Xiao, Milgram, & Doyle, 1997) or during realistic simulations (e.g., Manser, et al., 2007; Mosier, et al., 1998; Orasanu & Fischer, 1992). Observational techniques include field notes that researchers take during their observations, and (or) video- or audio-recordings of operator actions and interactions that are subsequently analyzed. While observational approaches present researchers with a wealth of information, both data collection and analysis pose formidable challenges.

Taking notes concurrently with observation is taxing to working memory and consequently may be susceptible to errors. Post-observational notes also may be affected by memory limitations potentially resulting in omissions or biased recall. Audio- or video-recordings, in contrast, provide an objective account of operators’ behavior, and support detailed performance analyses; however, at times at substantial costs (e.g., the study of team communication as well as the analysis of an individual’s think aloud protocol frequently are based on the transcription of all verbal behavior). In recent years, software programs have been developed that allow users to tag behavioral classifications (i.e. codes) to video-recordings, thus eliminating the need for labor-intensive transcription. We will review here several current tools (for more comprehensive reviews see Rohlfing et al., 2006, and Soria, et al., 2002) that have been widely used in behavioral research: two non-commercial tools, ANVIL (Kipp 2001; 2003), ELAN (Hulsbosch, van Uytvanck, & Hellwig, 2009), and two commercial ones, the FIT-system (Held & Manser, 2005), and the Observer XT (Noldus, 1991; Noldus, et al, 2000). While these tools provide considerable advantages over note-taking (such as increased reliability; time-stamped observations linked to behavior), they, too, come with limitations, most notably concerning their usability and flexibility.

The tools generally support user-defined codes; however they differ in the extent to which their interface design is amenable to a wide range of research questions and areas. Systems like ANVIL (Annotation of Video and Spoken Language) and ELAN (EUDICO Linguistic Annotator) have been specifically designed to facilitate the classification of human language behavior (e.g., linguistic and paralinguistic features; speech acts; gestures; non-verbal communication). The Observer XT and the FIT-system (Flexible Interface Technique), in contrast, have been used successfully to examine a variety of research issues in diverse domains, including usability testing (Kaikkonen, et al., 2005), nurse-patient communication (Uitterhoeve, 2009), team communication and performance (Guerlain, et al. 2005; Held & Manser, 2005), or ergonomic assessment (Held, Bruesch, Krueger, & Pasch, 1999).

Data entry in most systems is rather cumbersome, and typically requires users to pause the video. In ELAN and ANVIL users type their coding into the annotation board, or select a code from a pull-down menu of previously defined categories. Users of the Observer XT define coding categories prior to an observation and subsequently click on the appropriate category in the coding window. Codes can also be mapped onto keyboard shortcuts to speed up the coding process. However, remembering the correct code-to-key mappings may burden working memory, especially when multiple coding passes are performed on the same video-recording. A more user-friendly and less memory-taxing data entry solution is offered by the FIT-system. Here, users define their own interface by drawing or writing on a transparency overlay, and enter their codes by tapping the appropriate symbol on the overlay with a stylus.

An important conceptual distinction in observational research is between events as points occurring in time versus events as states extending over time. For instance, in the analysis of team problem solving, one might want to examine which team member contributes what information (= event-based analysis). Alternatively, one might want to characterize different problem solving phases in team members deliberations (= state-based analysis). ANVIL, ELAN, and the FIT-system do not account for this distinction. In contrast, users of the Observer XT can specify whether a code refers to the occurrence of an event (called 'point events') or to a state (called 'state events'). Events and states are subsequently represented in the Observer's visualization module as points and horizontal bars, respectively.

In cognitive engineering and NDM research, observations frequently concern multiple agents. It is therefore critical that observational tools support the cross-classification of behavior and agents. This facility is absent in ANVIL and ELAN, or could only be accomplished by characterizing the behavior of different agents in separate coding layers. To conduct agent-specific coding in the FIT-system, users need to perform two taps with the stylus, one on the icon of the agent, and the other on the icon of the behavior, presumably in a set order. This coding procedure may increase the observer's cognitive workload considerably and introduce errors as the number of agents and their possible actions increases, especially since the tool is used for real-time coding. A similar criticism may apply to the keyboard shortcut solution offered by the Observer XT insofar as each agent-behavior pairing is assigned a different key. The Observer's alternative procedure in which agent-behavior pairings need to be selected from a matrix in the coding window, could get unwieldy with larger numbers of agents and behavior types.

A final limitation of the tools reviewed is their focus on either video-annotation or coding of live performance. The ideal coding tool for cognitive engineers and NDM researchers should provide time-stamped classifications for both types of behavioral data. Neither ANVIL, ELAN or the FIT-system fulfill this requirement. ANVIL and ELAN will not support the classification of live behavior. The FIT-system, in contrast, has no video-tagging capability. The Observer XT was originally designed as a coding tool for the analysis of video data; however recently, a mobile version has been added.

ECOVRD—A NEW TOOL FOR EVENT CODING AND THE VISUAL REPRESENTATION OF DATA

The objective of our project was to create an observational tool that builds on the strength of existing technology but overcomes some of their limitations. We developed a tablet-based tool that enables users to perform real-time coding of behavior either from video-recordings or by observing live performance. The tool was designed both to advance behavioral research and to support applied uses, for instance in professional coaching. Several design goals motivated our work: (1) The tool should be easy to use, with the interface facilitating the coding process rather than adding to raters' workload; (2) It should support the coding of video-recorded and live behavior; (3) It should be flexible and not limited to specific research questions, allowing users to define coding categories; (4) Its coding feature should enable users to distinguish between events and states; (5) It should support the simultaneous assessment of multiple agents acting in concert or concurrently; and (6) It should provide users with an interactive visualization of observational data.

The current prototype of the tool was designed to support the analysis of video-recorded team interactions that occurred during computer-simulated search missions set in Antarctica (Fischer, McDonnell, & Orasanu, 2007). Teams (shown in screenshots in subsequent sections) consisted of four distributed members identified by different colors (Red, Blue, Green and Purple). Analyses focus on the cognitive and social aspects of team communication such as task coordination and interpersonal affect.

Salient Design & Interaction Features

Fluid User-Tool Interactions. ECOVRD has been specifically designed for use on a Tablet PC exploiting the new facilities that touch-based interfaces provide. The user has the freedom to rely on stylus, mouse or fingers as the input device. This feature allows for a much more fluid interaction between the user and the system, and supports our novel coding technique of tapping a category in a pie-menu (see description below).

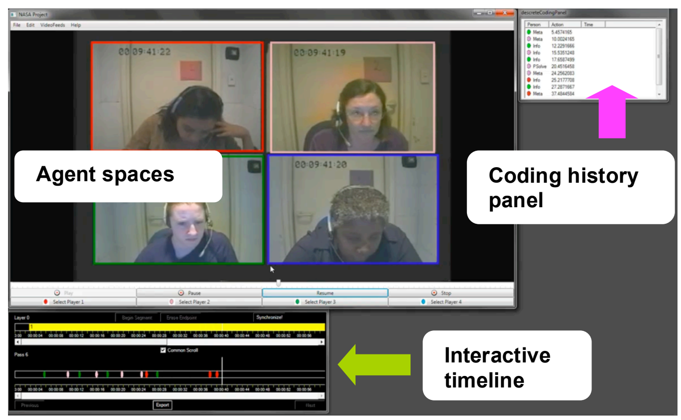

Figure 1. Interface layout of the coding module

Coding as a pointing action. The unique feature of our coding tool is that each agent is "assigned" an individual space on the input screen (see Figure 1). This design feature is consistent with human language processes –and thus should be intuitive to users—as it is derived from sign languages. Signers place the sign referring to a person or object in space and convey pronominal reference to this person or object by pointing to the appropriate location (Emmorey & Falgier, 2004). Analogously in our tool, coding categories will be overlaid as pie-menus on each space. Consequently, when a code is selected, it will be linked with the "agent occupying the space."

This space-based coding style affords users to link ratings to observed agents by performing one fluid movement towards a single location, rather than having to do multiple touches/clicks in different locations of the screen. As shown in Figure 2, tapping the code "Info" in the space of team member Green will be recorded as *"Speaker: Green; Action: Info Sharing"* where the user has configured the coding categories to be types of problem solving talk (Info Sharing; Problem Solving; Team Coordination; Meta-Cognition). Our pointing-based coding technique supports the simultaneous classification of actions by up to six agents. This number reflects screen real-estate considerations and is within the processing capacity of human working memory (Miller, 1956). Note that an agent space could be 'inhabited' by an individual or a team, for instance when observations concern the interactions between different departments or sub-teams in a company.

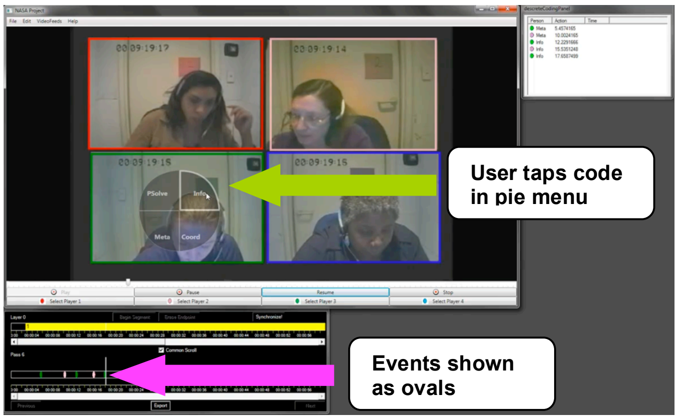

Figure 2. Event-based coding using pie-menu

Coding live performance or video-recorded behavior. Pointing-based coding supports real-time assessment of both live as well as video-recorded behavior. Figures 2 and 3 illustrate the input screen used to code interactions of a distributed team. Recordings of individual team members play in the background of their designated spaces outlined by the players' corresponding color.

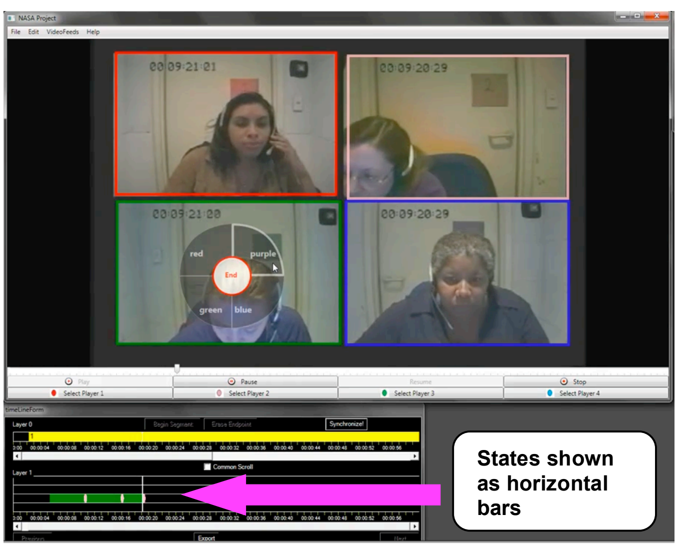

Figure 3. State-based coding using pie-menu

For co-located teams –i.e. when there is one video recording—the recording may be shown on a separate computer screen, or it may be displayed in the center of the tablet screen with the team member "spaces" surrounding it (the location of these spaces can be customized consistent with the user's needs). When users code live behavior or annotate recordings of co-located teams, team members' spaces will only be delineated on the screen (i.e., no individual videos playing in the background of agent spaces) and identified either by color or some other symbol.

Flexible coding categories. Users can custom-define the coding categories that appear in the pie-menu. This feature enhances the usability of our tool in many domains and for different research questions. In the tool's current version the number of coding categories is limited to seven consistent with the processing limits of working memory (Miller, 1956); especially, since our emphasis is on real-time classification. Users can also specify the nature of the coding categories as states—focusing on the duration of behavior (= state-based or interval coding)—or as discrete instances that occur over time (= event-based or discrete coding). Figure 2 shows the interface when users engage in discrete coding. Figure 3 illustrates interval coding. To initiate interval coding, users need to tap *Start* (which will then change to display *End*) and then tap on the coding category. In the example shown, coding captures which players interact with one another. As can be seen, Green initiated an exchange which involves at the given moment only Purple. To indicate the end of the interval, users need to tap *End*. The tool also has a note-pad feature which enables users to enter observations as text (Figure 4). Notes are time-stamped, tagged to the video (insofar as observations concern video-recorded behavior), and their existence is indicated with a symbol above the timeline.

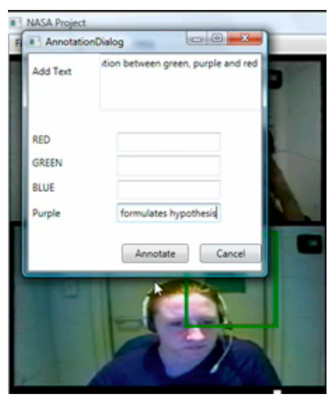

Figure 4. Note-pad to add comments

Intuitive workspace elements. Several interface elements further enhance the usability of our tool: (a) *Layered timeline* – Figure 2 shows a timeline in the bottom half of the interface. Its colored oval shapes correspond to observations a coder noted for each of the four team members. To represent multiple iterations of coding of the same video recording, layers were introduced on the timeline; thus each layer concerns a different aspect of team members' behavior. This feature is similar to the musical score format present in ANVIL and ELAN. In these systems as in our tool, each new coding layer is added below previous ones and entries at each layer run from left to right. This representation nicely captures the sequential occurrence of events within a layer as well as the temporal relation between events across different layers. Layers may be independent of each other, or may be in a parent-child relationship (i.e., a later coding pass builds on a previous one). When an observation in the parent layer gets refined by coding in the child layer, its node on the timeline is highlighted (brightened) to represent visually that the current coding is related to it. This display solution provides users with a quick overview of the coding process and enables them to re-visit coded segments since codes are time-stamped and tagged to the video. (b) *Editing feature* – Users can identify and isolate relevant segments in a video recording for subsequent coding. In particular, users can group coded observations that share relevant features and may occur within a coding layer or across different layers into a "bucket." The content of a bucket may then be displayed in a new coding layer for additional analysis or transferred to the visualization module for data summary. (c) *Coding history panel* – On the right side of the interface shown in Figure 2 is the coding history panel which provides users with an up-to-date account of recently created codes. Its simple grid-based design and the fact that agents are identified by different colors, make it easy for users to review their coding and to verify its accuracy.

Visualization module. This module presents users with quick overview options to 'eye-ball' the data. It also allows users to export the data into other data analysis tools, such as SPSS. As shown in Figure 5, the layered timeline appears on top, and a user can select agents that should appear in the visualization results by tapping on their respective color in the filter window. The user can then invoke a menu to create different visualization widgets which include options such as *Histogram, Pie-Chart, Radar-Map*.

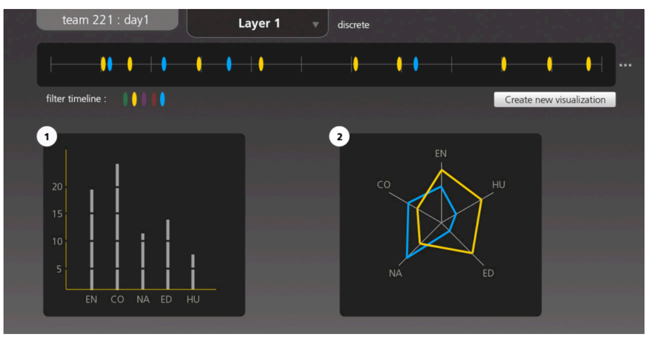

Figure 5. Two visualization widgets created by the user appear side-by-side in the visualization module of the tool

CONCLUSIONS AND FUTURE DEVELOPMENTS

ECOVRD has a user-friendly and intuitive interface and enables users to code in real time live behavior or video-recorded performance. The tool is highly versatile, as coding categories are user-defined and can capture instances of behavior or the duration of behavior. Thus, we anticipate that our tool will be useful both to the research community and in applied settings. For instance, ECOVRD could support flight instructors during the simulator training of flight deck crews. Currently instructors need to write down their observations while they are running the simulation. Our tool would not only facilitate instructors' assessments but would also provide them with time-stamped observations they could use during the debriefing session with pilots to quickly locate relevant behavior in the video-recording of pilots' performance. Similarly, our tool—in particular the information provided by the visualization module—could be valuable in team training. Graphic data summaries and basic statistics provide immediate performance feedback. Stored summaries of previous analyses of a team can be called up for comparison with the current analysis. This feature should be useful for observers and teams alike to detect salient changes in their interactions over time (e.g., decreasing participation of one team member).

Our next step will be to test our current prototype with researchers who pursue rather different questions than team communication analysis. We are also exploring the use of symbols, apart from colors, to distinguish between agents as this representation may be better suited for other research domains. In addition we are developing different options to present the coding categories. One possibility is to give users the option to select a grid format rather than a pie menu; this solution may be particularly attractive for users who wish to employ more than seven coding categories; for instance, Bales' 12 interpersonal process functions (Bales, 1976).

A further development will be to give users the flexibility to use different coding categories for different agents. Currently coding categories are identical across agents. In future application, user should be able to tie coding categories to specific areas (agents) and thus be able to classify concurrently occurring and related but different behavior by agents.

ACKNOWLEDGMENTS

Support for this research has been provided by NASA Cooperative Agreement NNX08AW90A to Ute Fischer. The authors thank three anonymous reviewers for their comments. Demo videos illustrating the features and use of ECOVRD can be viewed at http://ecovrd.lcc.gatech.edu .

REFERENCES

Badke-Schaub, P. (in press). Decision making processes of leaders in product development. In K. Mosier & U. Fischer (eds.), *Informed by knowledge: Expert performance in complex situations*. New York, NY: Taylor & Francis.

Bales, R. F. (1976). *Interaction process analysis: A method for the study of small groups (Revised)*. Chicago, IL: The University of Chicago Press.

Durso, F. T., Crutchfield, J. M., & Harvey, C. M. (2007). The cooperative shift change: an illustration using air traffic control. *Theoretical Issues in Ergonomics Science*, *8(3)*, 213-232.

Emmorey, K. & Falgier, B. (2004). Conceptual locations and pronominal reference in American Sign Language. *Journal of Psycholinguistic Research*, *33(4)*, 321-331.

Fischer, U., McDonnell, L., & Orasanu, J. (2007). Linguistic correlates of team performance: Toward a tool for monitoring team functioning during space missions. *Aviation, Space and Environmental Medicine, 78*(5), II, B86-95.

Guerlain, S., et al. (2005). Assessing team performance in the operating room: Development and use of a "black-box" recorder and other tools for the intraoperative environment. *Journal of the American College of Surgeons*, *200*(1), 29-37.

Held, J., Bruesch, M., Krueger, H. & Pasch, T. H. (1999). The FIT-System: A new hand-held computer tool for ergonomic assessment. *Medical & Biological Engineering & Computing*, *Supplement 2*(37), 862-863.

Held, J., & Manser, T. (2005). A PDA-based system for online recording and analysis of concurrent events in complex behavioral processes. *Behavior Research Methods*, *37*(1), 155-164.

Hellwig, B., van Uytvanck, D., & Hulsbosch, M. (2009). *ELAN: Linguistic Annotator, version 3.7*. Retrieved on January 30, 2009 from http://www.lat-mpi.eu/tools/elan/

Kaikkonen, A., Kekäläinen, A., Cankar, M., Kallio, T., & Kankainen, A. (2005). Usability testing of mobile applications: A comparison between laboratory and field testing. *Journal of Usability Studies*, *1*, 4-16.

Kipp, M. (2001). Anvil: A generic annotation tool for multimodal dialogue. In *Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech)*, 1367-1370.

Kipp, M. (2003). ANVIL 4.0 User Manual. Retrieved on January 5, 2009 from http://www.dfki.de/~kipp/anvil

Lauche, K. & Bayerl, P. S. (in press). Planning, monitoring, and trouble-shooting: Decision making in distributed drilling operations. In K. Mosier and U. Fischer (eds.), *Informed by knowledge: Expert performance in complex situations*. New York, NY: Francis & Taylor.

Manser, T., Harrison, T. K., Howard, S. K., & Gaba, D. M. (2007). Coordination patterns and clinical performance levels in the management of a simulated anesthetic crisis. In *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*, 658-662.

Miller, G. A. 1956. The magical number seven, plus or minus two: Some limits of our capacity to process information. *Psychological Review, 63*, 81-97.

Mosier, K. L., Skitka, L. J., Heers, S., & Burdick, M. D. (1998). Automation bias: Decision making and performance in high-tech cockpits. *International Journal of Aviation Psychology, 8*, 47-63.

Noldus, L. P. J. J. (1991). The Observer: A software system for collection and analysis of observational data. *Behavior Research Methods, Instruments, & Computers*, *23*(3), 415-429.

Noldus, L. P.J. J., Trienes, R. J. H., Hendriksen, A. H. M., Jansen H., & Jansen, R. G. (2000). The Observer Video-Pro: New software for the collection, management, and presentation of time-structured data from videotapes and digital media files. *Behavior Research Methods, Instruments, & Computers*, **32**, 197-206.

Orasanu, J., & Fischer, U. (1992). Distributed cognition in the cockpit: Linguistic control of shared problem solving. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp.189-194). Hillsdale, NJ: Erlbaum.

Patterson, E. S., & Woods, D. D. (2001). Shift changes, updates, and the on-call architecture in space shuttle mission control. *Computer Supported Cooperative Work*, *10*, 317-346.

Rohlfing, K., Loehr, D., Duncan, S., Brown, A., Franklin, A. Kimbara, I., et al. (2006). Comparison of multimodal annotation tools – a workshop report. *Gesprächsforschung Online-Zeitschrift zur verbalen Interaktion,* 99-123.

Schraagen, J. M., Leijenhorst, H. (1997). Searching for evidence: Knowledge and search strategies used by forensic experts. In E. Salas & G. Klein (eds.), *Linking expertise and naturalistic decision making* (pp. 263-274). Mahwah, NJ: Erlbaum.

Soria, C., Bernsen, N. O., Cadée, N., Carletta, J., Dybkjær, L., Evert, S., et al. (2002). Advanced Tools for the Study of Natural Interactivity. In *Proceedings of the Third International Conference on Language Resources and Evaluation* (LREC'02), 357-363.

Uitterhoeve, R., Bensing, J., Dilven, E., Donders, R., deMulder, P., Achterberg, T. van (2009). Nurse-patient communication in cancer care: Does responding to patient's cues predict patient satisfaction with communication. *Psycho-Oncology*, *18*(10), 1060-1068.

Zsambok, C. E. (1997). Naturalistic decision making: Where are we now? In C. E. Zsambok & G. Klein (eds.), *Naturalistic decision making* (pp. 3-16). Mahwah, NJ: Erlbaum.

Xiao, Y., Milgram, P., & Doyle, D. J. (1997). Capturing and modeling planning expertise in Anesthesiology: Results of a field study. In C. E. Zsambok & G. Klein (eds.), *Naturalistic decision making* (pp. 197-205). Mahwah, NJ: Erlbaum.