Systemic-Structural Activity Analysis of Video Data: A Practical Guide

Steven Robert Harris
Hypermedia Research Unit
University of Glamorgan
Wales, UK. CF37
+44 1443 483276
srharris@glam.ac.uk

ABSTRACT

Video data analysis can be used to support several aspects of the information technology (IT) design process, from initial requirements gathering, through prototyping and usability testing, to the development of support materials. This paper describes methods for the morphological (structural) and functional analysis of video data that involve the isolation and classification of tasks; the isolation and classification of actions; and the application of self-regulation models to the functional analysis of goal-directed activity during task performance. These methods have been developed within the context of a long-term research project involving field studies of collaborative computer-mediated activity by non-professional users in an education setting. Their theoretical foundation is in the systemic-structural theory of activity (SSTA), a modern synthesis of recent research in applied activity theory (AT) which is specifically tailored to the study and design of various aspects of human work activity.

Keywords

Video analysis, IT-design, activity theory, morphological & functional analysis, field data, systemic-structural theory of activity.

INTRODUCTION

Activity-theoretical approaches to human-computer interaction (HCI) and information technology (IT) design are centrally concerned with the multidimensional analysis and design of computer artifacts in relation to the complex, historically developing, sociocultural and technical context of their actual or proposed use. Video analysis provides a valuable resource for these efforts. Skillfully used, video recording can capture sequential data showing both the broad context and fine detail of IT use in real-world and laboratory settings; the possibly of unobtrusive (Blomberg, 1993) or unattended (Bauersfeld & Halgren, 1996) recording can help minimize the impact of observation on the situation under study. Supplemented by other observational methods, video provides a basis for the development of detailed descriptions of both individual and collaborative computer-mediated work activity, allowing the repeated review of complex and fleeting events, and giving researchers and designers the opportunity to check and amend previous interpretations in the light of new data or analytic insights (Suchman & Trigg, 1991; Jordan & Henderson, 1995).

Bødker's seminal work on focus-shift analysis using video data (Bødker, 1996) presented an activity theory-based approach to structuring video analysis in human-computer interaction (HCI) research. This article presents methods that aim to complement and extend that work. Whereas Bødker's work was mainly grounded in the general activity theory of Leont'ev (Leont'ev, 1978, 1981) as interpreted and developed by the Scandinavian school of IT-design (Bertelsen & Bødker, 2003), the methods described here have been developed within the framework of the systemic-structural theory of activity (SSTA), a distinctive and modern activity-theoretical approach specifically oriented toward the study and design of work and learning. SSTA represents a synthesis of recent AT research into practical problems in ergonomics, engineering psychology, and education, and draws on elements within activity theory - such as the concepts of self-regulation and the unity of cognition and behavior - so far under-utilized by Western researchers (Bedny & Meister, 1997; Bedny, 1998; Bedny et al., 2000; Bedny, 2004).

Scope of the paper

This paper focuses on the practical application of SSTA to video data analysis, addressing stages of video analysis following initial data capture, logging and archiving:

- The selection of data for analysis
- Transcription & the development of general time structures
- The isolation & classification of tasks
- Video data as a basis for morphological analysis
- Video data as a basis for functional analysis

The analytical techniques described may be applied sequentially, concurrently, or iteratively depending on the purpose and nature of the study.

SCOPE, BACKGROUND, & GROUNDING IN AT

Scope, Applications, and Outcomes of the Methods

The methods described here may be used to support those aspects of the IT-design process where the use of video data is appropriate, e.g. requirements gathering, cooperative prototyping, development of support materials, usability testing and general evaluation. They are suitable for application by researchers, analysts, and designers with a background in task analysis, and require some familiarity with the basic terms and principles of activity theory.

The immediate outcomes of the methods are an integrated set of descriptions - in textual, symbolic and diagrammatic forms, and at varying levels of detail and abstraction - of the structure and function of activity during computer-mediated task performance as recorded in the source data. The purpose of these descriptions is to support analysts in developing a detailed understanding of the relationship between the computer artifacts in use, task goals and conditions, sociocultural, physical, and technical context, and users' actual and possible strategies of activity in the work process. Depending on the stage and nature of the IT-design process in which they are used, these descriptions may prove sufficient. However, by virtue of having been developed within the integrated framework of SSTA, they may also be used as a basis for further analysis and modeling using other techniques of systemic-structural analysis, which include the development of detailed models of the temporal relationships between motor and cognitive actions during task performance, and the calculation of quantitative measures of task complexity.

Empirical Background

The techniques described in this paper have been developed in the context of a long-term participatory action research project involving the creative and collaborative use of information technologies by non-professional people from low-income, low education backgrounds (Harris & Shelswell, 2001; Harris, 2002, 2004; Shelswell, 2004). A principal concern of the research is to investigate how various aspects of IT-design either inhibit or support the development of fluent interaction in computer-mediated work and learning activity. To date, the project has included four phases of fieldwork, carried out in the setting of an Adult Basic Education (ABE) center in the South Wales Valleys region of the UK. The illustrative examples in this paper are drawn from the first and second phases of fieldwork - a longitudinal study in 2000-2001 followed by a shorter study in 2002 - where adult literacy and numeracy learners were observed as they used IT in collaborative media projects.

In the longitudinal field study, 27 subjects (14 male, 13 female, aged 15-73), took part in 103 sessions of collaborative computer-mediated activity in which they conceived, planned, and carried out individual and group projects in Web and multimedia authoring, digital video, computer graphics and animation, virtual reality (VR), and computer programming. During this phase of study the primary focus was on developing general, qualitative descriptions of collaborative activity using macro-analytical techniques. Data capture was mainly through participant observation and ethnographic interview. Single-camera video recording was used to opportunistically record instances of interaction breakdown and fluency.

The second phase of field research focused on the detailed analysis of activity during the performance of a single task, and used video recording as the primary means of data capture. One tutor and 8 learners (2 male, 6 female, aged 37 -76) were observed as they worked over 3 hours to produce paper documents using desktop publishing software. Learners worked collaboratively in three subgroups, each subgroup making use of 1 personal computer (PC). Task requirements were set by the tutor, with learners able to choose from a range of topics and formats for their documents within the stated constraints. A total of 6 video data sources were used: 3 tripod-mounted DV cameras, in conjunction with screen recording software running on each PC in use.

Grounding in Activity Theory

The Systemic-Structural Theory of Activity (SSTA)

SSTA formulates human activity as a logically ordered system of mental and behavioural actions, where cognition, behavior and motivation are integrated and organized by mechanisms of self-regulation toward achieving conscious goals (Bedny & Harris, 2004). Activity is understood as multidimensional and structured, being composed of discrete, hierarchically organized elements and involving four general stages: goal formation, orientation, execution, and evaluation (Bedny & Karwowski, 2003a).

As activity unfolds, mechanisms of self-regulation allow subjects to continually reconsider and adjust their goals and behavior strategies in response to changing conditions. An important aspect of this self-regulatory process is the ongoing comparison of the goal (a conscious, more or less precise representation of the desired future result of activity) with the result, i.e. actual outcome, of activity. If the result of an activity does not coincide with the subject's goal, then she or he must reformulate their strategy for goal achievement, or reformulate the goal itself. In this dynamic process, evaluation of the result is affected by functional subsystems such as subjective assessments of task difficulty and significance, subjective standards of successful results, etc. (for a general model of the functional structure of self-regulation see Bedny & Meister, 1997 p. 77).

SSTA approaches activity as a complex object of study embodying multiple, distinct aspects, and insists that multiple approaches must be employed to describe any single episode of activity. The systemic-structural analysis and design of activity correspondingly makes use of multiple analytical methods and units of analysis, integrated within a recursive, multi-stage, multi-level framework (Bedny & Karwowski, 2003b). Broadly, the analytical process involves identifying the available means of work, tools and objects; their relationship with possible strategies of work activity; existing constraints on activity performance; social norms and rules; possible stages of object transformation; and changes in the structure of activity during skills acquisition.

Systemic-structural analysis and design

In systemic-structural analysis, the structural elements of activity - *activity, task, action, operation, function block,* along with a composite unit called *member of algorithm* – provide the basic analytical objects and units. Figure 1 is a schematic representation of objects of study and units of analysis used in the systemic-structural approach. From the continuous flow of human activity, activity during task performance is selected as a primary *object* of study. The principal *units* used for morphological (structural) analysis are actions and operations; in functional analyses, macro- and micro- function blocks provide the basic units.

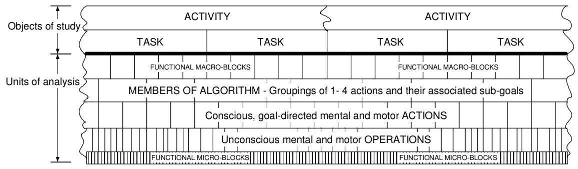

Figure 1. Objects of study and units of analysis in systemic-structural activity analysis.

The analysis and design of activity based on SSTA utilizes three distinct, but closely interrelated, approaches: the parametrical, which focuses on studying various parameters of activity using techniques such as error analyses and cognitive (process) analyses; the morphological, which focuses on the description of the structure of activity during task performance as a series of discrete actions and operations; and the functional, which describes activity as a goal-directed, self-regulating process. All three approaches are involved in some or all of four general stages of analysis: (1) qualitative description, (2) algorithmic analysis, (3) time structure analysis, and (4) quantitative (complexity) analysis. These four stages are recursively related; later stages of analysis usually require revisiting earlier descriptions in order to refine or refocus them. Each stage can be carried out from the point of view of some or all of the three approaches and at different levels of detail or decomposition; depending on the requirements of the research, not all stages may be appropriate or required.

Other SSTA approaches include objectively logical analysis, which integrates several methods including verbal and graphical description of tasks, description of technological processes, tools, equipment, conditions of work; sociocultural studies of the context in which task performance takes place; and the study of individual-psychological factors such as features of personality and individual style of activity performance. This paper mainly focuses on stage 1 and 2 analyses from the individual-psychological perspective using the morphological and functional approaches, and begins by discussing the selection of data for analysis.

SELECTION OF DATA FOR ANALYSIS

When using time-based audiovisual data, analysts need to strike an appropriate balance between in-depth analysis and the needs and constraints of a particular project. Analysis time to sequence time ratios (AT:ST) typically range from 5:1 to 100:1, while detailed micro-analyses may approach ratios as large as 1000:1 (Fisher & Sanderson, 1996). In selecting data for analysis, the research projects described here followed the general outlines given in Trigg, Bødker, & Grønbaek, (1991), Suchman & Trigg (1991), and Bødker (1996). In accord with Engeström's approach to work development and IT-design research (Engeström, 1991; Engeström & Escalante, 1996; Engeström, 2000), Bødker recommends that analysts use interaction breakdowns and focus shifts as indicators of potentially design-relevant episodes in the data. In systemic-structural terms breakdowns can be defined as forced changes of subjects' strategies of action caused by unacceptable divergence between actual results and action or task-goal. In relation to IT-design, breakdowns suggest that subjects' existing approaches to problem solution with the application are proving inappropriate, and that opportunities for learning-in-use may be inadequately supported by the artifacts in use or the use context.

In the longitudinal study described above, which took an opportunistic approach to video capture, interaction breakdown was selected as a primary trigger both for initial recording and subsequent analysis. A paper-based instrument, the Breakdowns Pro-Forma (BDPF), was designed to support the identification and preliminary qualitative analysis of interaction breakdown and recovery across all longitudinal study data, including video. The BDPF could be completed concurrently with, or immediately following, observation or recording; during compilation of the field note record; retrospectively during field note review; and during initial review and logging of the video data. Figure 2 shows an example of a BDPF completed during initial video data review. Pro-forma categories included a unique identifier, date, time, task details, the application(s) in use, nature of the problem(s), interface elements involved, whether and how recovery was achieved, user and researcher comments, and codes for the relevant video segment and related data sources. BDPF data was subsequently compiled into a spreadsheet and linked to the database of video and interview transcripts.

In contrast to the longitudinal study, the short field study was specifically designed to capture all interaction in as much detail as possible. In this case the resulting body of video data was logged and transcribed in its entirety prior to further analysis. Following the identification of task and sub-task sequences within the recorded activity, functional analyses were conducted on the video data and sets of transcripts arranged into sequences representing activity by each participant sub-group (see below). Incidents of interaction breakdown and notably fluent interaction were then used as a basis for selecting data for detailed morphological analysis. Functional analysis was applied to the whole dataset.

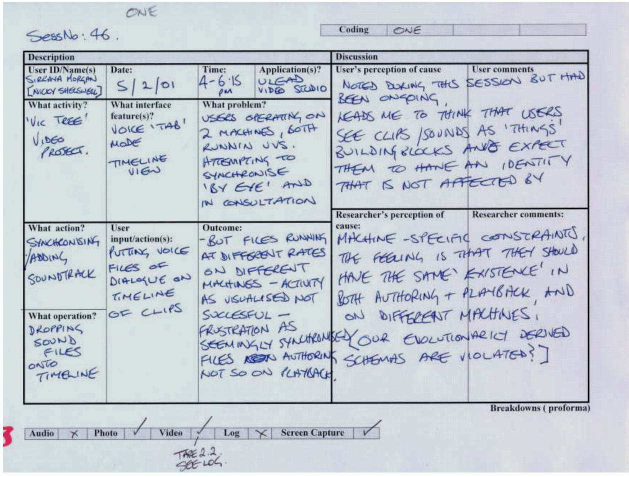

Figure 2. Completed Breakdowns Pro-Forma.

TRANSCRIPTION & DEVELOPMENT OF GENERAL TIME STRUCTURES

The production of accurate transcriptions of subjects' speech actions during task performance is central to the approach described here, where time-coded transcripts provide a basis for other analytical techniques, in addition to being an important resource in their own right[1]. Transcription is not, however, simply about producing transcripts. Engaging in transcription engenders the close familiarity with observed events that is one of the principal advantages of working with video data; it is recommended that, wherever practicable, persons performing activity analyses should also be responsible for transcribing the materials. The alternating focus between audio and video sources, scanning both verbal and motor behavior, encourages awareness of their complex interconnectedness, and of the interlocking individual-psychological and sociocultural aspects of use activity.

There is now a wide range of software tools available to facilitate video transcription and annotation (a useful review can be found in Sanderson et al., 1994). In the studies discussed here, transcription was carried out using Transana[2]. This is an open-source video transcription and analysis tool that supports frame-by-frame viewing and the accurate linking and playback of transcripts and video segments. Transcript lines can be automatically numbered; linking frames in a video to positions in the transcript is achieved through the embedding of time codes into the text files. Transana will automatically highlight the relevant portion of the transcript during video playback. The tool also supports various coding and searching operations (Thorn, 2002, see Figure 3.).

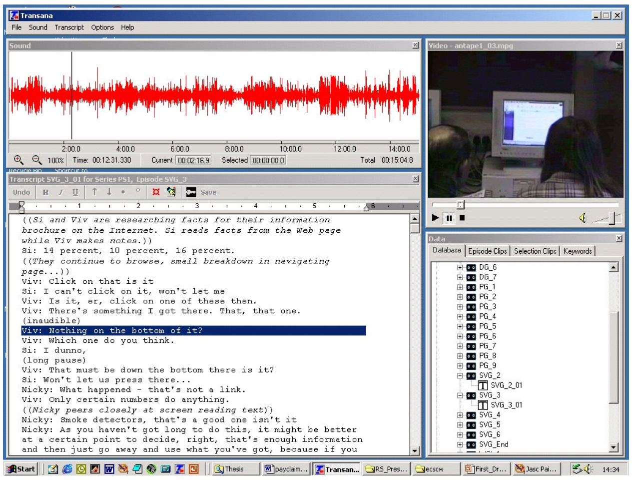

Figure 3. Transana interface, showing Waveform, Transcript, Video and Database windows.

Compiling Transcript Sequences

For practical reasons such as capacity of storage media, multiple camera sources, etc. video data is usually captured, stored and transcribed in discrete segments, often of varying size. For systemic-structural video analyses such as in the short field study described above, it is useful to compile separate segment transcripts into sequences representing the whole period of activity of interest for the study. Where multiple cameras and other video sources are used, multiple transcript sequences can be used to represent a whole observation session from the varying points of view of any of the sub-groups or individuals in the study. Similarly, in some cases it can be appropriate to use an editing application to compile video segments into longer sequences showing the whole arc of development of a particular task activity.

ISOLATION & CLASSIFICATION OF TASKS

The notion of activity during task performance as a fundamental object of study underpins systemic-structural video analysis. In SSTA, a task is understood a sequence of goal-directed actions involving an initial situation (the problem presented before task performance begins), a transformational situation (actions taken to solve the problem), and a final situation (initial situation changed). Video analysis can be used to study how the structure of a task changes during different stages of performance, and to identify how many basic transformational stages are required.

Tasks are organized around a supervening goal, with the vector motive→goal determining the directedness of activity during task performance (Leont'ev, 1978). Tasks are always carried out under specific circumstances - task

---

[1] In common with other AT approaches SSTA grounds dialogic analysis of intersubjective interactions in the work of Bakhtin. See Bedny and Karwowski (2004a p. 139).

[2] Transana was developed by the Wisconsin Center for Education Research and is available for free download from www.transana.org.

conditions - which determine the constraints on performance. These include interacting situational elements, rules, and alternatives for situation transformation. Tasks are structured by requirements that help to specify the goal, such as instructions or commands, and are also affected by the means of work in the given conditions, and the raw materials or input information being explored or transformed. In some circumstances task requirements and conditions may contradict each other, causing breakdowns in activity. Task attributes include complexity, subjective assessments of difficulty, and significance. These attributes involve elements (such as subjects' past experience) which cannot be directly evinced from video data but must be established through other techniques such as interviews or historical analysis. This emphasizes the need to integrate video analysis with other methods of research. Bødker (1996) and Kaptelinin *et al* (1999) have proposed the use of checklists in support of this integration process, an approach adopted by the field studies reported here although not directly discussed in this text.

Isolation of Tasks

Transcript sequences provide a basis for identifying the logical structure and temporal sequence of task and sub-task solution from the ongoing flow of activity. These general mappings underpin task classification and are a precursor to the techniques of morphological analysis described in the following section. In functional analyses, they allow comparison of the structure of actual task performance with objectively set task conditions, e. g. they enable the comparison of observed task-solving strategies with those given in verbal instructions, manuals, etc., and as presented in subjects' verbal protocols. Using line numbers as reference points, sequential transcripts are sub-divided into segments, based on identification of the task or subtask in which participants are engaged.

The upper limit at which the isolation of tasks begins, and which sequences of activity are considered as primary or sub-tasks, is study-dependent, and will also be practically determined by the nature of the video data. For example, in analyzing video from the short field study, a broad differentiation was made between the initial portion of the 3-hour session, where the activity of the whole group was guided by the goal setting task conditions based on instructions by the tutor; the majority of the session, when sub-groups were involved in task solution; and a final part, when the whole group was involved in giving feedback on their outcomes. Each of these sections was then further sub-divided into sequences of activity organized around distinctive task-goals. This decomposition continued until the level of discrete actions[1] was reached, producing a 4-level hierarchical representation of sequentially ordered tasks and sub-tasks. Figure 4 shows a schematic representation of this task structure arranged against an approximate "timeline" formed by taking counts of line numbers in the transcript sequence. When required, more accurate time-structure diagrams of activity at the task level can be constructed using video time codes or other methods of time measurement as markers.

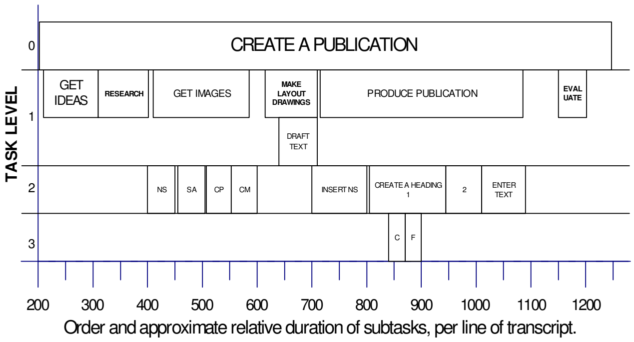

Figure 4. Structure & sequence of tasks in short field study.

As noted above, tasks and sub-tasks are primarily differentiated on the basis of their organizing goals. The formation or acceptance of a distinctive task-goal is taken as marking the inception of a task or sub-task. Achievement or abandonment of a goal is taken to mark task completion. Identification of these junctures is achieved through close examination of the video and transcript. Goal formation or acceptance is often associated with some orienting activity; in collaborative work this may be observed as discussion, sketching, note-taking and so on, as in the transcript extract below:

227 S: So what we gonna do then?
228 S: Health and safety sounds alright doesn't it?
229 T: Health & safety in the home? Ok, so now you want to jot down all the (inaudible)
230 S: (to V) Why has (inaudible) or something like that?
231 (T talks to other students while S and V discuss the topic. Mostly inaudible.)
232 (V is drawing on a pad on the table)

The next phase of a task is marked by sequences of executive actions organized around the task goal, e.g.:

887 ((S scrolls through fonts menu in Publisher))
888 T: If you want something a bit fancy there's Victorian, I dunno what that
889 S: ((turns to V)) That's quite nice innit?
890 V: Yes
891 S: Want that?
892 V: Alright
893 T: Or you can, you can bold it if you want... and then, if you want a shape... ((points to menu))
894 S clicks to pull down shapes menu))
895 T: then you can choose whether you want it to go round, or wiggly, or slanting, or any of those
896 S: which one do you want V?
897 V: Anything you fancy
898 V: I just gonna say (something like)
899 S: Tha one or tha one?
890 ((indicates with cursor))
891 V: Now then, try that one is it, yeah

[1] By using both the notion of task and action, SSTA more clearly defines the differing structural levels of activity which have been presented in a confusing and contradictory manner in some of the ATIT literature, e.g. (Kuutti, 1996, Figure 2.4).

Followed by evaluation:

985 S: Eh, that's nice innit?

986 T: Not very readable though, that's the only trouble with that one

Figure 5 shows the decomposition of activity during the short field study from the point of view of one sub-group, using transcript line numbers as a reference.

Task/Subtask name	Subtask level	Begin	End	Duration
Create a publication	0	201	1248	1047
Get ideas	1	223	291	68
Research	1	293	393	100
Get images	1	410	613	203
Get image of no-smoking sign	2	407	458	51
Get image of smoke alarm	2	463	482	19
Get image of chip pan	2	504	554	50
Get image of child w/matches	2	556	594	38
Make layout drawings	1	624	707	83
Draft text	1	645	707	62
Produce publication	1	712	1098	386
Insert no smoking picture	2	712	802	90
Create a heading 1	2	810	962	152
Create a heading 2	2	962	1001	39
Choose heading color	3	837	865	28
Choose font	3	866	893	27
Enter text	2	1002	1094	92
Evaluate project (V only)	1	1158	1187	29
Collect copies	0	1274	1276	2

Figure 5. Task decomposition from video transcript.

Classification of Tasks

Once tasks have been identified it is then possible to classify them according to various criteria. Distinctions between task types are relative rather than absolute, and are based on assessing the varying degrees of freedom of performance associated with individual tasks. Those tasks which are highly structured by the artifacts in use (e.g. data entry) may be considered as predominantly deterministic, or deterministic-algorithmic; that is, they require a standardized sequence of actions for successful completion. Probabilistic-algorithmic tasks are those which require choices at some stages. Depending on the interaction between task requirement and conditions, such tasks often become problem-solving tasks for the subject. In cases where task uncertainty is even greater, tasks are termed heuristic or semi-heuristic, the major criterion in these categorizations being the extent to which the task presents an undefined field of solution. Probabilistic-algorithmic tasks often also include non-algorithmic problem-solving components; in the same way, semi-heuristic task-problems may also include algorithmic and semi-algorithmic sub-tasks (Landa, 1983; Bedny & Harris, 2004).

It is important to re-emphasize at this point that the methods described here can only model some aspects of the full complexity of task performance; the more unstructured the activity, the more vital it is that they are used in conjunction with other techniques. Classification of tasks will clearly be more or less relevant depending on the activity under study and the stage and nature of the IT-design process. Task typing may prove useful during initial project scoping, as it can be significantly more difficult to design work processes that involve a high levels of freedom of performance, creativity, and unpredictability. In requirements gathering, task classification can underpin initial design specifications. In evaluation, applications, prototypes, or designs may be assessed against the types of tasks in which they are expected to be, or are actually, used.

Video analysis supports the study of when and how subjective perceptions of the nature of the task lead subjects to adopt unexpected or inappropriate strategies of action, and the influence of aspects of the design of computer artifacts in use on this process. Video can also be useful for identifying when there is a mismatch between the type of task an application is designed for and its actual uses. For example, in the longitudinal study it was found that while the video editing application in use employed a sequential metaphor and modal functionality, supporting a step-by-step approach to task performance, most of the tasks for which it was used could be classified as semi-heuristic, and required much greater flexibility from the application interface. Noting this divergence between design and use at the task level provided a useful way of thinking about the large number of interaction breakdowns experienced with this application (Harris, 2004).

VIDEO DATA AS A BASIS FOR MORPHOLOGICAL ANALYSIS

Comparative analysis of the structure of activity with the physical and logical configuration of the equipment that mediates action and provides the conditions under which activity is performed provides a useful basis for IT-design. A fundamental principle of activity theory is the unity of cognition and behaviour (Bedny, Karwowski, & Bedny, 2001). The structure of activity during task performance is understood as formed from a logically organized system of external-behavioral and internal-mental actions and operations. This structure continually develops in response to internal and external conditions, and involves both conscious and unconscious self-regulation mechanisms. This section discusses the morphological analysis of video data, using action as a primary analytical unit, and focusing on the interconnectedness of, and transition between, mental and motor actions. It describes the morphological analysis of video data through (1) the isolation and classification of discrete actions, and (2) the generation of algorithmic descriptions of the logical structure of activity.

Definition of Action

An action is defined as a relatively complete element of activity that fulfills an intermediate, conscious goal. Actions are temporal: the initiation of a conscious goal (goal acceptance or goal formulation) constitutes their starting point; they conclude when the actual result of action is evaluated in relation to the goal. This understanding of action allows the systematic description of the continual flow of activity during task performance as divided into individual units. Actions can be described in terms of a non-linear, recursive loop structure, with multiple feed-forward and feedback interconnections. Figure 6 shows a highly simplified model of goal-oriented action as a one-loop system.

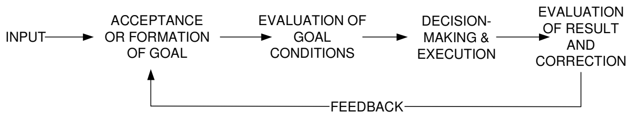

Figure 6. Action as a one-loop system.

Isolation of Actions

In order to isolate discrete actions it is necessary to identify the goal, object, and tools involved in *each* action. The nature of an action is dependent on the interrelation of these components in any particular situation. A useful approach to isolating individual actions in a sequence of video-recorded task activity is to begin by setting out a basic sequential description of the technical steps involved. Figure 7 shows the description of a Web authoring task recorded during the longitudinal study. Using such a description as a way of orienting toward the data, analysis can then proceed to the isolation of actions. For example, Table 1 shows the isolation of actions involved in Step 2c of Figure 7, "in the display pane, move cursor to hyperlink being checked, left-click on link" and illustrates their classification (according to two typologies outlined in the following section). It can be seen that some actions involve several tools. Where tools are not defined (as in action 1) this indicates motor activity not involving external instruments – although AT always assumes that motor actions contain cognitive components – and may involve the use of "internal" psychological tools. Such tools can be assumed in action 2, which implicates not only the perception of signs visible on the interface but also the use of concepts and images to interpret them.

Table 1. Example of task isolation & classification.

Action	*Goal*	*Object*	*Tools*	*Type*
1	Reach & grasp mouse	mouse		Object-practical
2	Locate display pane	Interface	Graphical interface elements	Sign-practical/ Simultaneous -perceptual
3	Move cursor over hyperlink	cursor	mouse	Object-practical
4	Activate link	hyperlink	Cursor, left mouse button	Sign-practical

It should be emphasized that in systemic-structural theory, the concept of "tool" is tightly associated with the concept of action; outside of a specific task, it is often not possible to precisely determine whether material or ideal artifacts are acting as tools mediating a specific action, or as objects of that action. This suggests that personal computers, or software packages, cannot simply be classed as "tools" of activity except when conducting broad macro-analyses. Rather, they should be considered as means of work that present (or create) a variety of material and symbolic objects that, at different stages during the performance of computer-based tasks, may either mediate actions as tools, be the object of actions, or simply provide the conditions under which actions are performed (cf. Bødker, 1991, 1999).

1. Launch or restore focus to Web browser (IE)
   a. If application already running move cursor onto window area and left click *or* move cursor onto taskbar icon and left click.
   b. If application not running, move cursor to application icon and left-click *or* move cursor to start button, left-click, navigate to appropriate menu, choose application icon or label and double left-click.
2. Load or refresh appropriate HTML document
   a. If document is already being displayed, move cursor to refresh icon (or select command from View pull-down menu) and left-click
   b. If document not displayed, load into browser by either selecting File>Open from pull-down menu, then typing file path or browsing to file location in Open dialogue box *or* use left mouse button and cursor to drag file icon from desktop or other location and drop on browser display pane by releasing mouse button
   c. In the display pane, move cursor to hyperlink being checked, left-click on link
3. View resulting display and assess
4. Make decision on whether link needs to be removed or altered
5. Launch or restore focus to text editor (Notepad)
   a. If application already running move cursor onto window area and left click *or* move cursor onto taskbar icon and left click.
   b. If application not running move cursor to application icon and left-click *or* move to start button, left-click, navigate to appropriate menu, choose application icon or label and double left-click.
6. Load or review appropriate text (.html) file
7. Identify section of markup that corresponds to link being checked
8. Cut, delete, or modify link markup
9. If moving location of link:
   a. Identify new location in appropriate text file
   b. Navigate text editor to appropriate file and location
   c. Insert cursor and paste link markup
10. Return to Step 1.

Figure 7. Sequence of basic technological procedures in the task "Update Web Pages".

Classification of Actions

A number of different approaches to the classification of actions have been developed in system-structural theory. Two were used in the studies reported here. The first differentiates types of cognitive, or mental actions based on two considerations: (1) the degree to which they require deliberate examination and analysis of the stimulus (direct connection with or transformation of the input) and (2) their dominating psychological process during performance: sensory, simultaneous perceptual, imaginative, mnemonic, etc. The second classification scheme is more generalized, categorizing actions according to the nature of their object, which may be either material or a sign or symbol, and according to their method of performance, either practical (motor) or mental. This scheme distinguishes *object-practical actions* performed with material objects, *object-mental actions* performed on mental images, *sign-practical actions* performed with external signs, and *sign-mental actions* performed through the mental manipulation of signs or symbols. Table 1 provides examples of classification using both schemes. When required, standardized motor actions may also be identified and categorized using established measured time and motion systems such as MTM-1. This approach can be helpful in determining the interrelationship of mental and motor actions (in Bødker's terms, analyzing the impact of the *physical* aspects of the computer artifacts in use on the structure of activity) and when using time-structure analysis (see Figure 9 and discussion below).

Video Data as a basis for Algorithmic Analysis

The identification and classification of actions during task performance provides a basis for developing models of the logical structure of activity, using symbolic representations known as human algorithms. Human algorithms make use of an additional unit of analysis, called *member of algorithm*. These are formed from clusters of 3-5 actions, organized by a supervening goal. The construction of algorithms supports detailed examination of the performed actions and logical relationships in a task or sub-task. A completed algorithm consists of specialized notation accompanied by explanatory text. The symbols called *operators* denote clusters of efferent (O^{ϵ}) and afferent (O^{α}) mental and motor actions; (*l*) the probabilistic logical conditions that structure their relationships; with arrows (↓↑) and sub- and superscript numbers indicating the various logical links between operators and logical conditions. Exhaustive explanations of this syntax and examples of algorithmic analysis are beyond the scope of this paper but can be found in (Bedny et al., 2000; Bedny, Karwowski, & Kwon, 2001; Bedny & Karwowski, 2003b and ; Sengupta & Jeng, 2003). It is important to note the algorithmic modeling of activity produces idealized, abstract models of activity to be used in conjunction with the various other methods described here and elsewhere in the systemic-structural literature; these techniques should not be taken to imply that the SSTA view of human activity considers it deterministic or rigidly tied to plans.

In the recursive process of systemic-structural video analysis, the development of human algorithms encourages reviewing of the video and other data in a systematic manner, bringing additional insights. In the longitudinal study, algorithms were used to support the identification of how, and at what point, aspects of tool design contributed to interaction breakdown. Figure 8 shows a fragment from the algorithmic description of the Web authoring task discussed earlier. Here, the subject was using multiple applications (text editor, browser, file manager), each with different functionality and interface features. Construction of the algorithm helped to identify those points in the task where complexity was maximal, both highlighting aspects of the video for closer scrutiny and raising issues not readily apparent from the data. In this case, this led to the identification of conflicts between the various ways the applications in use handled windowing as the major contributor to interaction breakdown.

Member of algorithm	*Description of Members of Human Algorithm*
$\overset{7}{\downarrow}\overset{11(1)}{\downarrow}O^{\alpha}_{1}$	Identify appropriate Web browser window
$l_1\overset{1}{\uparrow}$	If correct Web browser window has focus go to O^{α}_{3}; if not go to O^{ϵ}_{2}.
O^{ϵ}_{2}	Bring browser window to foreground by moving cursor onto window area and left-clicking *or* moving cursor onto taskbar icon and left-clicking.
$\overset{1}{\downarrow}O^{\alpha}_{3}$	Check to see if correct Web page is displayed
$l_2\overset{2}{\uparrow}$	If appropriate Web page is displayed go to O^{ϵ}_{4} ; if not go to O^{ϵ}_{5}
O^{ϵ}_{4}	Refresh browser display by moving mouse cursor to refresh icon and left-clicking *or* selecting command from View pull-down menu and left-clicking
$\overset{2}{\downarrow}O^{\epsilon}_{5}$	Open correct HTML document by either selecting File→Open from pull-down menu, and browsing to file location using Open dialogue box *or* use left mouse button and cursor to drag file icon from desktop or other location and drop on browser display pane by releasing mouse button
$\overset{4}{\downarrow}\overset{4}{\downarrow}O^{\alpha}_{6}$	Identify next hyperlink to be checked
O^{ϵ}_{7}	Activate link by moving mouse cursor to hyperlink being checked and left-clicking
$O^{\alpha\, th}_{8}$	Look at new page display in browser window and compare with expected result
$l_3\overset{3}{\uparrow}$	If browser display is appropriate go to O^{ϵ}_{9} ; if "The page cannot be displayed" HTML file is displayed *or* if link works but page inappropriate go to O^{ϵ}_{11}

Figure 8. Fragment of algorithmic description of activity during a Web authoring task.

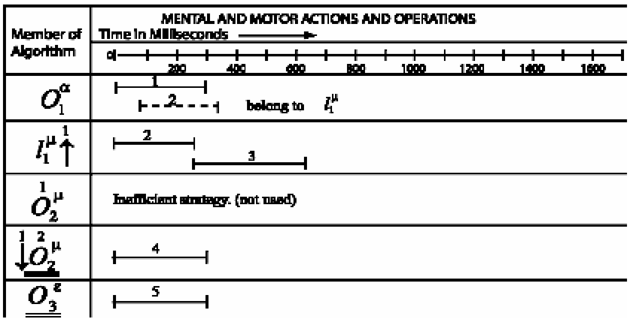

Figure 9. Fragment of time structure of computer graphics task. From Sengupta & Jeng (2003).

Having mapped the logical structure of task performance also makes it possible to describe the temporal structure of activity in terms of performed actions, supplementing the timeline charts used for task analysis as described earlier. Time measurements derived from the video data can be used to specify the duration of individual elements of activity. Attention is paid to the structure of sequential and simultaneous performance of mental and motor actions. Figure 9 shows a fragment from the time-structure of an HCI task using a word-processing application developed in a lab-based study by Sengupta & Jeng (2003). This method can be extended through use of standardized classifications of motor actions from e.g. MTM-1 (see Bedny & Meister, 1997 pp. 252-262).

VIDEO DATA AS A BASIS FOR FUNCTIONAL ANALYSIS

As noted above, SSTA understands activity as a psychologically self-regulating system, where the structure of activity is changed by factors such as the subject's changing goals and experience. Functional analyses trace various aspects of this self-regulation process, at different levels of detail and different stages of activity. They produce functional descriptions of cognitive action which, firstly in conjunction with other methods of individual-psychological analysis such as parametrical and morphological analyses, and then with social-historical and objectively-logical descriptions, can be used to produce a holistic understanding of activity (Bedny & Karwowski, 2004b). Functional analysis of video data proceeds by using self-regulation models as "lenses" with which to "scan" the activity captured on video tape. Self-regulation models depict functionally invariant components (function blocks) whose content changes as activity unfolds, and which mutually affect each other through feed-forward and feedback influences. Function blocks may be used for the analysis of observed activity at various levels of detail (see Figure 2); video data produced by ethnographic field studies will most often be appropriate for analysis using functional macro-blocks. These blocks represent functional entities, such as the goal, which are understood as "black boxes" comprising coordinated systems of subfunctions.

Functional analysis of video data can be used to study issues such as how users' goals, motivation, and experience regulate their interaction with the systems in use. Activity-theoretical IT-design moves beyond traditional cognitive approaches by understanding that perception and sense-making are active processes critically affected by the goal of activity; according to the goal of subjects' activity, interface elements may have no significance or not even be perceived at all (Bertelsen, 2003). As a cognitive mechanism, the goal can be more or less clear and precise, and either totally or partly conscious. For example, in the type of explorative activity which can constitute a significant part of IT use – especially for less experienced users - the goal may be initially vague, gradually becoming more precisely formulated as the object becomes clearer.

In such explorative activity, goal-formation is closely associated with the process in which, as a subject attempts to solve a task-problem, he or she creates a level of aspiration that may be changed by evaluation of the results of activity. As a subject performs exploratory actions, they produce and analyze a sequence of trials-and-errors, leading to the creation of a hypothesis about the situation, and the formulation of a preliminary goal for activity. This preliminary goal can be considered as corresponding to a particular level of aspiration, inasmuch as it is tied up with evaluations of task difficulty that includes not only assessment of the objective characteristics of a task, but also elements of self-evaluation. In the research project in which these methods have been used, which is concerned with IT use by individuals from low-income, low-education backgrounds, it has been of fundamental interest to find methods of assessing how the design of computer artifacts in use affects this process. Support for goal-formation has emerged as a critical factor in the development of fluent technology use by study participants.

Using Video Data to Trace the Goal Formation Process

In the short field study, functional analysis of the video data was used to investigate how the computer artifacts in use supported goal formation. The example described here focuses on the collaborative computer-mediated activity of two subjects, S and V, as they undertake the loosely defined desk-top publishing task. Successful completion of the task is dependent on their formation of a "shared vision" (a sufficiently aligned task-goal) of the form and content of the publication. This requires the continual adjustment of their individual goal-images and strategies for action, through dialogue and examination of actual outcomes on the screen of the PC in use. The observed work involved extensive explorative activity investigating specific functions of the DTP software, and the search and retrieval of images and information from the Internet, all framed within ongoing discussion. Once sufficient precision and alignment of their goal-images had been achieved, a task-problem solution adequate to the requirements was quickly achieved.

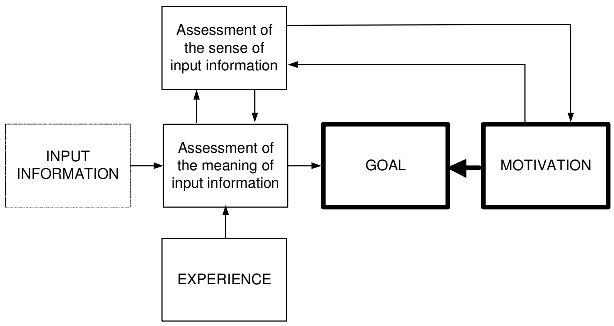

Figure 10. Functional model of goal formation. *After Bedny (1997).*

Figure 10 shows the self-regulation model used in the functional analysis. This simplified model[1] includes (macro) function blocks representing goal, motivation, experience, and the assessment of sense and meaning of input information in the activity under analysis. Analysis proceeded through systematic examination of the video data, transcripts, and other evidence in order to trace the changing contents of the different function blocks. In the

[1] Examples of more detailed models of self-regulation are presented in Bedny & Meister (1997 p. 77), and Bedny & Karwowski (2003, 2004b).

example here, it can be seen that input information was generated by the graphical display presented by the (shared) computer monitor, verbal instructions from the tutor, and the collaborating subjects' dialogue; these sources were studied using the video footage and transcripts. Data on long-term aspects of the subjects' experience and motivation (individual-psychological profiles, social background, previous observed behavior, etc.) were derived from sources such as interviews and attendance and achievement records. Assessments of how subjects' experience, motivation and interpretation of the sense and meaning of input information developed during task performance came from examining the record of motor and verbal actions.

The Time Structure Outline

In support of the process of functional analysis a tabular form based on the transcript sequences was developed. The Time Structure Outline (TSO) correlates durations (indicated by transcript line ranges) with brief descriptions of task-related events depicted by the video and transcript data. A fragment of one TSO is shown in Figure 11.

Line Nos.	Description
291-292	V signals end of 'get ideas' subtask, T confirms.
293-393	Sub-task 'research'. PC becomes major means of work, Web browser, Search Engine (text search) main material tools mediating activity. S becomes main tool operative. List artifact created by V also supports activity.
297-300	S encounters some physical/behavioral difficulty in assuming the role as PC operator due to physical environmental (space, configuration) and individual (ambivalent handedness) constraints.
367-379	Breakdown in interaction with Web Browser as S & V wish to use text item on Web page perceived as hyperlink that does not respond to left mouse-click. This slows hitherto fluent interaction with site, as users become unsure how to read the semantic properties of the page. Goal-directed action becomes more exploratory, as users pursue what the page can give them rather than what they purposively seek – page becomes object of actions. On return to group, T confirms non-hyperlink status of (underlined?) text and activity moves forward again.
382-392	T negotiates cessation of sub-task 'research', although S & V have not indicated that they feel the goal has been achieved – the subtask goal is too imprecise to clearly guide activity. The tutor, T, is regulating the activity from the point of view of the task goal 'project to be completed in one 2 hour session'.
393	V confirms willingness to end sub-task 'research'

Figure 11. Fragment of time-structure outline.

The TSO provides a cross-cutting view into the dataset, facilitating the systematic interrogation of the data from the point of view of each of the function blocks in the self-regulation model in use. As shown in Figure 11, the table can also be graphically annotated as an aid to interpretation, using e.g. shading and simple symbol systems. In the example here annotations indicate incidences of breakdown, and mark passages of interest with regard to goals, tool use, and usability issues. Shading indicates the start and end of subtasks, facilitating the integration of functional analysis with the methods described in previous sections.

During the functional analysis of S and V's task-solving activity, the TSO helped to identify and describe various stages of the goal-formation process as evidenced in the data. For example, the TSO entry below:

> S repeatedly passes her hands across the screen as she speaks, indicating (to herself & V) where the block elements of the page layouts should go. The image-goal is now precise enough for her to implement it directly; at 704, 705, 706, S & V exchange utterances that indicate that they are now both visualizing a similar image & text.

Refers to the following transcript extract:

691 S: Erm, that smoking one might fit if you have it up on
692 like on trains and buses, so what would be good for that
693 one is erm, smoking, smoking over there.
694 V: See if I can get this 26 percent
695 S: Ah. Ah you could put that right... 26 percent ...of
696 people in their house, right, cause this is the house
697 one innit - that's right - then you could put erm
698 ((Gestures across screen without speaking to
699 indicate layout and text))
700 ((indicator of goal-image precision))
701 S: 26 percent of people in the house..and
702 V: How about if you put "smoking kills 26 percent"
703 S: In houses
704 V: Smoking in bed (S: ah yes) when a cigarette..
705 S: Falls out the ashtray
706 V: and it goes down the side of a chair
707 S: Go on then.

The relevant frames of video (e.g. Figure 12) show the collaborators leaning in toward the screen, and S using gestures to indicate the position of text and image elements in the document being created. Subsequent video segments and screen recordings show the process of attempting to place screen elements in the indicated positions. Using these, transcripts, and a copy of the artifact produced by the activity, it becomes possible to study aspects of how the application in use supported achievement of the task-goal.

Figure 12. Collaborative production of a DTP document.

CONCLUSION

This paper has briefly outlined activity theory-based methods that can be used in support of the IT-design process. They involve the generation of structural and functional descriptions of human work activity based on video data analysis. The primary object of study is activity during task performance, making use of analytical units – action, member of algorithm, function block – as developed and defined within the systemic-structural theory of activity. The aim of the methods is to develop understandings of the interrelationship between the structure of work activity, its functioning, and the configuration of the ideal and material components of work. SSTA emphasizes that activity is multi-faceted and must be studied from a variety of mutually supporting perspectives. This paper has concentrated on presenting techniques that take an individual-psychological perspective on work analysis, as this level of description is currently under-developed in ATIT. At this level of analysis, collaborative work is examined mainly from the point of view of each individual subject. In order to develop a fuller picture, the resulting models must then be coordinated with each other and with descriptions developed from the perspectives of socio-cultural-historical and objectively-logical analysis. Another paper in this workshop reflects on the methods described, their relationship to other approaches in activity-theoretical IT-design, and the practical experience of applying them in a design-oriented research project.

ACKNOWLEDGMENTS

Thanks to staff and students at Coleg Morgannwg Open Learning Centre, Pontypridd, Wales, UK for their participation in the research reported here; to Tirthankar Sengupta & One-Jang Jeng for kind permission to reproduce the illustration used in Figure 8; and to the reviewers of earlier drafts of this paper for their helpful comments.

REFERENCES

Bauersfeld, K., & Halgren, S. (1996). "You've got three Days!" Case Studies in Field techniques for the Time-Challenged. In D. Wixon & J. Ramey (Eds.), *Field Methods Casebook for Software Design.* New York: John Wiley & Sons.

Bedny, G. Z. (1998). Activity Theory. In W. Karwowski (Ed.), *International Encyclopedia of Ergonomics and Human Factors*. New York: Taylor & Francis.

Bedny, G. Z. (2004). Preface, special issue on activity theory. *Theoretical Issues in Ergonomics Science, 5*(4), 249-254.

Bedny, G. Z., & Harris, S. R. (2004). The Systemic-Structural Theory of Activity: Applications to the Study of Human Work. To appear in *Mind, Culture and Activity, 12*(2).

Bedny, G. Z., & Karwowski, W. (2003a). Functional Analysis of Orienting Activity and Study of Human Performance. *In Proc. Ergonomics in the Digital Age, IEA Congress 2003* (Vol. 6, pp. 443-446). Seoul: The Ergonomics Society of Korea.

Bedny, G. Z., & Karwowski, W. (2003b). A Systemic-Structural Activity Approach to the Design of Human-Computer Interaction Tasks. *International Journal of Human-Computer Interaction, 16*(2), 235-260.

Bedny, G. Z., & Karwowski, W. (2004a). Activity theory as a basis for the study of work. *Ergonomics, 47*(2), 134-153.

Bedny, G. Z., & Karwowski, W. (2004b). A functional model of human orienting activity. *Theoretical Issues in Ergonomics Science, 5*(4), 255-274.

Bedny, G. Z., Karwowski, W., & Bedny, M. (2001). The Principle of Unity of Cognition and Behavior: Implications of Activity Theory for the Study of Human Work. *International Journal of Cognitive Ergonomics, 5*(4), 401-420.

Bedny, G. Z., Karwowski, W., & Kwon, Y.-G. (2001). A Methodology for Systemic-Structural Analysis and Design of Manual-Based Manufacturing Operations. *Human Factors and Ergonomics in Manufacturing, 11*(3), 233-253.

Bedny, G. Z., & Meister, D. (1997). *The Russian Theory of Activity: Current Applications to Design and Learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bedny, G. Z., Seglin, M. H., & Meister, D. (2000). Activity theory: history, research and application. *Theoretical Issues in Ergonomics Science, 1*(2), 168-206.

Bertelsen, O. W. (2003). Activity Walkthrough: a cognitive walkthrough in activity theory terms. In M. Hertzum & S. Heilesen (Eds.), *Writings in Computer Science: Proceedings of the Third Danish HCI Research Symposium* (Vol. 98, pp. 17-20). Roskilde, Denmark: Roskilde University.

Bertelsen, O. W., & Bødker, S. (2003). Activity Theory. In J. M. Carroll (Ed.), *HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science* (pp. 291-234). London: Morgan Kaufmann.

Blomberg, J. (1993). Ethnographic Field Methods and Their Relation to Design. In D. Schuler & A. Namioka (Eds.), *Participatory Design: Principles and Practices* (pp. 123-157). Hillsdale: Lawrence Erlbaum Associates.

Bødker, S. (1991). *Through the Interface: A Human Activity Approach to User Interface Design*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bødker, S. (1996). Applying Activity Theory to Video Analysis: How to make Sense of Video Data in Human-Computer Interaction. In B. M. Nardi (Ed.), *Context and Consciousness: Activity Theory and Human-Computer Interaction* (pp. 147-174). Cambridge, MA: MIT Press.

Bødker, S. (1999). *Computer Applications as mediators of design and use - a developmental perspective.* Unpublished Doctoral Dissertation, Aarhus University, Aarhus.

Engeström, Y. (1991). *Developmental Work Research: Reconstructing expertise through expansive learning*. Amsterdam: Elsevier.

Engeström, Y. (2000). Activity Theory as a Framework for Analyzing and Redesigning Work. *Ergonomics, 43*(7), 960-974.

Engeström, Y., & Escalante, V. (1996). Mundane Tool or Object of Affection? The Rise and Fall of the Postal Buddy. In B. M. Nardi (Ed.), *Context and Consciousness: Activity Theory and Human Computer Interaction* (pp. 325-373). Cambridge, Mass: MIT Press.

Fisher, C., & Sanderson, P. M. (1996). Exploratory sequential data analysis: exploring continuous observational data. *interactions, 3*(2), 25-34.

Harris, S. R. (2002). PD in Ponty: Design-by-Doing in Adult Basic Education. In T. Binder & J. Gregory & I. Wagner (Eds.), *PDC 02 Proceedings of the Participatory Design Conference* (pp. 278-283). Malmö, Sweden: CPSR.

Harris, S. R. (2004). *Design, Development, and Fluency: An Activity-Theoretical Approach to Human-Computer Interaction Research.* Unpublished PhD. Thesis, University of Glamorgan, Wales, UK.

Harris, S. R., & Shelswell, N. (2001). Building bridges across the digital divide: supporting the development of technological fluency in Adult Basic Education learners. In A. Mason (Ed.), *Proceedings of the FACE Annual Conference 2001* (pp. 42-51). Southampton: FACE.

Jordan, B., & Henderson, A. (1995). Interaction Analysis: Foundations and Practice. *The Journal of the Learning Sciences, 4*(1), 39-103.

Kaptelinin, V., Nardi, B. M., & Macaulay, C. (1999). The Activity Checklist: A Tool for Representing the "Space" of Context. *interactions, 6*(4), 27-39.

Kuutti, K. (1996). Activity Theory as a potential framework for human-computer interaction research. In B. M. Nardi (Ed.), *Context and Consciousness: Activity Theory and Human Computer Interaction* (pp. 17-44). Cambridge, Mass: MIT Press.

Landa, L. M. (1983). Descriptive and Prescriptive Theory of Learning and Instruction. In C. M. Reigeluth (Ed.), *Instructional Design, Theories, and Models: An Overview of their Current State* (pp. 55-74). Hillsdale, NJ: Lawrence Erlbaum Associates.

Leont'ev, A. N. (1978). *Activity, Consciousness and Personality* (M. J. Hall, Trans.). Englewood Cliffs, N. J.: Prentice-Hall.

Leont'ev, A. N. (1981). *Problems of the Development of the Mind.* Moscow: Progress.

Sanderson, P. M., Scott, J. J. P., Johnston, T., Mainzer, J., Watanabe, L. M., & James, J. M. (1994). MacSHAPA and the enterprise of Exploratory Sequential Data Analysis (ESDA). *International Journal of Human-Computer Studies, 41*(5), 633-681.

Sengupta, T., & Jeng, O.-J. (2003). Activity Based Analysis for a Drawing Task, *In Proc. Ergonomics in the Digital Age, IEA Congress 2003* (Vol. 6, pp. 455-458). Seoul: The Ergonomics Society of Korea.

Shelswell, N. (2004). *A Critical Investigation into Uses of New Technologies in the Teaching and Learning of Adult Literacy.* Unpublished MPhil Dissertation, University of Wales Institute, Cardiff, Cardiff.

Suchman, L. A., & Trigg, R. (1991). Understanding Practice: Video as a Medium for Reflection and Design. In J. Greenbaum & M. Kyng (Eds.), *Design at Work: Cooperative Design of Computer Systems*. New Jersey: Lawrence Erlbaum Associates.

Thorn, C. (2002). *Creating New histories of Learning for Math and Science Instruction: Using NVivo and Transana to manage and study large multimedia datasets.* Paper presented at the Conference on Strategies in Qualitative Research, Institute of Education, University of London.

Trigg, R., Bødker, S., & Grønbaek, K. (1991). Open-ended interaction in cooperative prototyping: A video-based analysis. *Scandinavian Journal of Information Systems, 3*, 63-86.

Systemic-Structural Activity Analysis of Video Data: Experience & Reflections

Steven Robert Harris
Hypermedia Research Unit
University of Glamorgan
Wales, UK. CF37
+44 1443 483276
srharris@glam.ac.uk

ABSTRACT

Another paper in this workshop presents a practical guide to video analysis using methods based on activity theory. This short paper reflects on the historical and theoretical background of those methods, their development, and the practical experience of their use in the context of a design-oriented research project.

Keywords

Video analysis, IT-design, activity theory, morphological & functional analysis, field data, systemic-structural theory of activity.

INTRODUCTION

In their recent review of the use of activity theory (AT) in human computer interaction (HCI) research, Bertelsen and Bødker (2003) stressed that one of the "really big challenges" for activity theory informed IT-design is:

> …how far one may actually be able to go? How close to technology? How design-oriented?

They go on to cite the Finnish developmental work-research tradition as "for the time being the most complete methodological approach to activity theoretical work analyses and design" while noting that very few researchers in that tradition have, as yet, addressed issues of IT in use and design, forcing practitioners interested in such issues to "seek their own approaches". They conclude by suggesting that two interesting issues for future AT-based IT-design are the development of design instruments and technical concepts.

In another paper submitted to this workshop (Harris, 2004), I describe some outcomes of attempts, by myself and my colleagues, to seek out approaches to applying AT to understanding IT-design problems. Those attempts have taken place in the context of an ongoing research project involving empirical HCI field studies of the use of information and communication technologies by participants in adult basic literacy and numeracy education in the UK. The main purpose of the project is to investigate how aspects of IT-design inhibit or support the development of fluent interaction in computer-mediated activity, especially with regard to people like our participants, who are predominantly from low-income, low-education backgrounds, and who are thought to be – by the UK government at least - somewhat representative of those on the other side of the "digital divide". The aim of the research is to develop detailed understanding of computer artifacts in use as a basis for informing the future design of inclusive work and learning technologies.

APPROPRIATING ACTIVITY THEORY

Adult Basic Education takes place in a complex cultural and political context, and involves teaching and learning practices that has been developing over several hundred years. The relatively recent introduction of new technologies into ABE (in the mid-to-late 1990s) has had a profound impact, with both negative and positive effects (Harris & Shelswell, 2001; Harris, 2003). Since the start of our investigations in 1997, we have been seeking to identify frameworks and methods to guide and enable our research efforts that are appropriate to both our research setting and our questions. Early in the project, activity theory emerged as a potentially useful theoretical framework for the research. The decision to attempt to make use of AT was influenced by our contact with the Scandinavian tradition in participatory design (PD) and computer-supported collaborative work (CSCW). The political and social stance of those approaches seemed relevant to the radical empowerment tradition within ABE, of which the learning centre where the bulk of our research has taken place is a notable representative. Through our investigations into, and attempts to apply, PD (see Harris, 2002) we came into contact with the "tool perspective" (Ehn & Kyng, 1987; Ehn, 1988) and the seminal work of Susanne Bødker, Kari Kuutti and others involved with the introduction of AT into computer technology use and design research (Bannon, 1991; Bødker, 1991; Kuutti, 1991).

We followed up these initial contacts by seeking out as much material on the practical applications of AT as we could find. This turned out be more difficult, and unsatisfactory, than anticipated. Our primary sources became the 1996 collection edited by Bonnie Nardi (Nardi, 1996b), and texts such as Engeström's work (Engeström, 1987; Engeström & Middleton, 1996; Engeström et al., 1999), the Wertsch collections (Wertsch, 1981; Wertsch 1985; Wertsch et al., 1995), and the two widely available books by Vygotsky (Vygotsky, 1978, 1986). Although we found much stimulating material, few articles seemed immediately applicable to our practical research needs. One notable exception was the seminal paper on focus-shift analysis by Bødker (Bødker, 1996), which became an important starting point for many of the methods we subsequently developed, and which extensively informs the approach to video analysis presented in this workshop.

For us, one striking aspect of this review was an apparent disjunction between the wide range of approaches represented in the 1981 Wertsch collection, which contains articles (e.g. Zinchenko & Gordon, 1981) describing activity theory methods in terms which connected with our understandings of cognitive science and cybernetics, and the articles in the Nardi collection which seemed to mainly make use of what has become known as CHAT (cultural-historical activity theory). The difficulties we encountered in interpreting the foundational texts by Leont'ev (Leont'ev, 1977, 1978, 1981) did little to clarify matters. What emerged from our (limited) readings in the literature were a number of issues where we clearly had to "seek our own way". These included the question of how to delimit the amount of context to take into account – what is an activity, where does an "activity system" begin and end? Another concern was the identification of appropriate units of analysis for the body of field data we were building up, and what methods we should use to apply them. What are the practical implications of using *activity* or *activity system* as a basic unit of analysis, when what you are attempting to analyze *is* activity?

DEVELOPING A SYSTEMS-STRUCTURAL APPROACH

A step toward beginning to resolve some of these issues came through contact with the work of G. Z. Bedny and his associates on the systemic-structural theory of activity (SSTA). In particular, Bedny's 1997 textbook *The Russian Theory of Activity: Current Applications to Design and Learning* (1997, with David Meister) and its companion piece *Activity theory: history, research and application* (Bedny et al., 2000) served to introduce us to new viewpoints on AT, via a large body of work - carried out in the former Soviet Union and other Eastern Bloc countries – which addresses the application of AT to practical problems in ergonomics, engineering psychology, education, vocational training, and other disciplines.

We began to investigate ways of using the array of methods presented in this work to the requirements of our research. So far, results have been encouraging; the methods of systemic-structural video analysis developed on this basis reported in the companion paper are proving effective in generating design-oriented insights into the interrelationships between the structure of work activity, and the configuration of the material components of work. Specifically, SSTA has helped us to resolve the issue of what to study, when to study it, and what units of analysis to use.

SSTA presents a multi-dimensional approach to activity analysis based on employing multiple approaches and units of analysis within an integrated framework that involves three major analytical perspectives: socio-cultural-historical, objectively-logical, and individual-psychological (Bedny & Karwowski, 2004b pp. 258-260). The methods described in the companion paper mostly take the individual-psychological perspective, where the research focuses on the individual structure of each subject's activity, and the relationships between the activity of various subjects. In order to produce a holistic understanding of activity, studies from this perspective clearly need to be related to understandings generated from the other two; however, as it is precisely the detailed study of activity from the individual-psychological perspective that is notably absent in many of the extant AT-IT studies, this approach has been emphasized here.

Figure 2 in the companion paper presents a general scheme of the structural components of activity, which can be summarized here as:

Activity → Task → Action → Operation → Function Block

From the individual-psychological perspective, the first two components of this schema (activity, task) are considered primarily as the *objects* to be studied in activity analyses, the remainder being the *units of analysis* employed for the study of those objects. Because activity is understood as a holistic, structured, systemic entity, when an activity becomes the object of study it must be approached from multiple perspectives, using a variety of interrelated units of analysis. In this way, the primary object of study, some specific activity, can become the subject of a number of methodologically different, but closely interrelated descriptions. Initially, these are based on qualitative and parametrical analyses of the general characteristics of the activity of interest, identifying the available means of work, tools and objects; their relationship with possible strategies of work activity; existing constraints on activity performance; social norms and rules; possible stages of object transformation; and changes in the structure of activity during skills acquisition. In subsequent stages of activity analysis it is the *task*, understood as some fragment of activity that is organized around a supervening goal, which becomes the central object of study; the analytical focus shifts to activity during the solution of some practical problem.

THE TASK AS AN OBJECT OF STUDY

In IT-design, the notion of task has been introduced from varying perspectives in ergonomics, psychology and software design, and is generally differentiated from related terms by its implication of intentionality and human significance – the task is what someone wants to accomplish through the use of a computer. However, despite the widespread use of task analysis methods in IT-design, difficulties in clearly defining and using the term within the limitations of cognitive psychology have led some to question its utility. Draper (1993) was among the first to point out the multiple meanings of task in traditional HCI research, and the dangers of design based on rigid task analysis. Recognition of the shortcomings of existing task analysis methods such as GOMS has led some researchers using AT to reject the notion of task as entirely unhelpful (e.g. Holland & Reeves, 1996; Nardi, 1996a). However, in SSTA the task is considered a primary object of study, being defined as a sequence of actions organized around a supervening goal, with the vector motive→goal determining the directedness of activity during task performance (Leont'ev, 1978). The identification, isolation and classification of discrete tasks provides a means of delimiting the object of study from the ongoing flow of activity, providing a basis for the application of analytical units (action, operation, function block) which can describe tool-mediated actions at very fine levels of detail while retaining the connection to the holistic and developing context.

As noted earlier, restoring the centrality of the task as an object of study also goes some way toward resolving an issue that now appears to be impeding the development of activity-theoretical IT-design: the question of defining and applying appropriate units of analysis. This problem becomes acute in the "middle ground" of activity description, as researchers try to describe the structure of activity in terms of actions and operations, and to delineate the differences and transitions between the conscious and unconscious, and external and internal aspects of subject- and object-oriented activity. The extremely complex nature of these interrelationships can easily lead researchers seeking useful simplifications to generate interpretations which do little but confuse the issue. For example, consider the schematic representation of levels of activity given in Kuutti (1996 Fig. 2.4, p.33), a widely cited article which provides a good example of the problems faced by those pioneering researchers who took on the responsibility of introducing AT into IT-design. Tasks involving multiple actions organized around supervening goals - such as "fixing the roofing" are presented at the action level, while "selecting appropriate wording" is considered as an unconscious operation, and "building a house" an activity. SSTA suggests that it can often be more useful to think of such examples in terms of *production processes* which involve sequences of tasks, which in turn are composed of chains of internal and external actions with both conscious and unconscious aspects (Bedny & Harris, 2004; Bedny & Karwowski, 2004a).

The situation where models of activity suitable for one kind or level of description prove inappropriate to the particular demands of other stages or aspects of activity analysis can also be seen in an increasing number of applications to problems in IT-design of the well-known triadic schema developed by Engeström (Engeström, 1987; Engeström & Escalante, 1996; Engeström, 2000). This simplified model, which by representing various elements of an "activity system" supports a number of useful ways of thinking about and generating high-level qualitative descriptions of the dynamic relationships between various aspects of activity, is often unsuitable for the detailed analysis of tasks and actions. In particular, by failing to clearly distinguish between the goals, objects, and actual results of activity, or between object- and subject-oriented aspects of activity, this schema can lead to confusing and conflicting findings when applied to the analysis of activity during task performance from the individual-psychological point of view. For example, the analysis in Engeström (2000 pp. 961-4, especially Figures 1 & 2) results in tasks (examining and diagnosing the patient) being portrayed as actions, subjects (patient, father, lung specialist) being portrayed as objects, and "the computer as a technological instrument" being, apparently, operated unconsciously.

USING VIDEO IN TASK ANALYSIS

A central concern of our research in Wales has been to understand how IT-design supports or hinders the development of fluent interaction. In accord with other findings, our analyses indicate that, as task complexity increases, users' sense-making and problem-formulation actions become increasingly critical (Byström & Järvelin, 1995). SSTA clearly differentiates between task complexity and task *difficulty*, considering complexity as an objective property of a task, and difficulty as the subject's evaluation of the effects of that complexity. This distinction points to how the design of software may often support under- or over-representation of task complexity, thereby influencing subjective evaluations of task difficulty in ways that can significantly affect task solution, and to how acceptance or formulation of a task-goal is closely associated with subjective representations of task difficulty. An application (such as the video editing package used in our longitudinal study) may initially attract users by representing the task as less complex than it actually is, subsequently proving ineffective in supporting task solution through being insufficiently helpful in supporting users to understand the actual complexity of the task in hand. Conversely, a tool may make a relatively simple task seem much more complex than it actually is, discouraging the user from attempting, or continuing to attempt a task solution. A notable example of this latter case in our studies has been the software volume control running on the PCs in use, which frequently caused breakdowns. Investigating the influence on users' task performance of the relationship between the objective presentation of task and its subjective representation has become of major interest as we have begun to apply the techniques of functional analysis described in the methods paper to our video data.

In our research video has become an invaluable resource for studying the complex and constantly developing structures and processes of computer-mediated activity during task performance. To paraphrase Bødker (1996 p. 148), we cannot bring the world to a standstill while we do our analysis; what we can do is use video technology to continually review past events in the world from varying perspectives until our understanding develops.

The approach to systemic-structural video analysis outlined in the companion paper is situated within what Sanderson and Fisher (1994) have described as the cognitive tradition of exploratory sequential data analysis (ESDA). We are making use of the framework of systemic-structural theory to connect the detailed analysis of verbal and motor actions in task performance with the broader sociocultural and physical context of interaction. In terms of Fisher & Sanderson's (1996) taxonomy of eight fundamental smoothing operations during the analysis of time-based data, the video analysis methods mainly concentrate on *coding* the data according to activity-theoretical concepts, and on *conversions* of the data through the use of the various representational devices developed within SSTA. It is important to note that there is no suggestion that video analysis should be carried out in isolation from the other aspects of activity analysis. In particular, the various types of qualitative analysis which are concerned with delineating the socio-cultural-historical and objectively-logical aspects of the observed activity must be kept in constant dialogue with the video data. Bødker's work on focus-shift analysis has used checklists to facilitate this dialogue, as have we; in the research reported here, we have also found the development of short, descriptive texts helpful.

CONCLUSION

In our experience, the methods of video analysis presented to this workshop begin to clarify ways in which the study of human work activity based on activity theory can be used to inform the IT-design process at the fine levels of detail required if we are to "get close to the technology". The combination of broad qualitative studies of the general features of activity with the algorithmic and time-structure analysis of the sequence of actions during task performance points to the development of sufficiently detailed models which retain their connection to the context and complexity of holistic activity. In our work on video analysis we have begun to develop various analytical instruments which appear to have potential for supporting the IT-design process. This process has largely, though not exclusively, been made possible by making use of the concepts and principles of the systemic-structural theory of activity, which seems to offer as complete a methodological approach to activity theoretical work analyses and design as any we have so far encountered. It is hoped that, by continuing to "seek our own approach" on this basis, our future work will include the further development of technical concepts appropriate to the specific concerns of IT-design.

ACKNOWLEDGMENTS

I would like to express my gratitude to K. B. Bærentsen for first drawing my attention to systemic-structural theory, and to G. Z. Bedny for comments on earlier drafts of these papers.

REFERENCES

Bannon, L. J. (1991). From human factors to human actors: The role of psychology and human-computer interaction studies in system design. In J. Greenbaum & M. Kyng (Eds.), *Design at*

Work: Cooperative Design of Computer Systems (pp. 25-44). Mahwah, NJ: Lawrence Erlbaum Associates.

Bedny, G. Z., & Harris, S. R. (2004). The Systemic-Structural Theory of Activity: Applications to the Study of Human Work. *To appear in Mind, Culture and Activity, 12*(2).

Bedny, G. Z., & Karwowski, W. (2004a). Activity theory as a basis for the study of work. *Ergonomics, 47*(2), 134-153.

Bedny, G. Z., & Karwowski, W. (2004b). A functional model of human orienting activity. *Theoretical Issues in Ergonomics Science, 5*(4), 255-274.

Bedny, G. Z., & Meister, D. (1997). *The Russian Theory of Activity: Current Applications to Design and Learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bedny, G. Z., Seglin, M. H., & Meister, D. (2000). Activity theory: history, research and application. *Theoretical Issues in Ergonomics Science, 1*(2), 168-206.

Bertelsen, O. W., & Bødker, S. (2003). Activity Theory. In J. M. Carroll (Ed.), *HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science* (pp. 291-234). London: Morgan Kaufmann.

Bødker, S. (1991). *Through the Interface: A Human Activity Approach to User Interface Design*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bødker, S. (1996). Applying Activity Theory to Video Analysis: How to make Sense of Video Data in Human-Computer Interaction. In B. M. Nardi (Ed.), *Context and Consciousness: Activity Theory and Human-Computer Interaction* (pp. 147-174). Cambridge, MA: MIT Press.

Byström, K., & Järvelin, K. (1995). Task Complexity Affects Information Seeking and Use. *Information Processing & Management, 31*(2), 191 - 213.

Draper, S. (1993). The notion of task. In S. Ashlund & K. Mullet & A. Henderson & E. Hollnagel & T. White (Eds.), *Bridges between Worlds, INTERCHI'93 Conference Proceedings* (pp. 207-208). Reading, MA.: Addison-Wesley.

Ehn, P. (1988). *Work-Oriented Design of Computer Artifacts*. Stockholm: Arbetslivscentrum.

Ehn, P., & Kyng, M. (1987). The collective resource approach to system design. In G. Bjerknes & P. Ehn & M. Kyng (Eds.), *Computers and Democracy - a Scandinavian Challenge* (pp. 17-57). Aldershot: Avebury.

Engeström, Y. (1987). *Learning by Expanding*. Helsinki: Orienta-Konsultit.

Engeström, Y. (2000). Activity Theory as a Framework for Analyzing and Redesigning Work. *Ergonomics, 43*(7), 960-974.

Engeström, Y., & Escalante, V. (1996). Mundane Tool or Object of Affection? The Rise and Fall of the Postal Buddy. In B. M. Nardi (Ed.), *Context and Consciousness: Activity Theory and Human Computer Interaction* (pp. 325-373). Cambridge, Mass: MIT Press.

Engeström, Y., & Middleton, D. (Eds.). (1996). *Cognition and Communication at Work*. Cambridge: Cambridge University Press.

Engeström, Y., Miettinen, R., & Punamäki, R.-L. (Eds.). (1999). *Perspectives on Activity Theory*. Cambridge: Cambridge University Press.

Fisher, C., & Sanderson, P. M. (1996). Exploratory sequential data analysis: exploring continuous observational data. *interactions, 3*(2), 25-34.

Harris, S. R. (2002). PD in Ponty: Design-by-Doing in Adult Basic Education. In T. Binder & J. Gregory & I. Wagner (Eds.), *PDC 02 Proceedings of the Participatory Design Conference* (pp. 278-283). Malmö, Sweden: CPSR.

Harris, S. R. (2003). *Motivation, Participation and Co-location: Key aspects of Technology-based learning in Adult Basic Education.* (CASE Technical Report CTR0041). Wales: University of Glamorgan. Available for download from http://case.glam.ac.uk/CASE/StaffPages/SteveHarris/SRH.html.

Harris, S. R. (2004). Systemic-Structural Activity Analysis of Video Data: A Practical Guide: Paper presented to the 1st International Workshop on Activity Theory Based Practical Methods for IT Design, Copenhagen, Sept. 2-3, 2004.

Harris, S. R., & Shelswell, N. (2001). Building bridges across the digital divide: supporting the development of technological fluency in Adult Basic Education learners. In A. Mason (Ed.), *Proceedings of the FACE Annual Conference 2001* (pp. 42-51). Southampton: FACE.

Holland, D., & Reeves, J. R. (1996). Activity Theory and the View from Somewhere: Team Perspectives on the Intellectual Work of Programming. In B. M. Nardi (Ed.), *Context and Consciousness: Activity Theory and Human-Computer Interaction* (pp. 257-281). Cambridge, Ma.: MIT Press.

Kuutti, K. (1991). Activity theory and Its Applications to Information Systems Research and Development. In H.-E. Nissen & H. K. Klein & R. Hirschheim (Eds.), *Information Systems Research: Contemporary Approaches & Emergent Traditions* (pp. 529-550). Amsterdam: North-Holland.

Kuutti, K. (1996). Activity Theory as a potential framework for human-computer interaction research. In B. M. Nardi (Ed.), *Context and Consciousness: Activity Theory and Human Computer Interaction* (pp. 17-44). Cambridge, Mass: MIT Press.

Leont'ev, A. N. (1977). Activity and Consciousness (R. Daglish, Trans.), *Philosophy in the USSR: Problems of Dialectical Materialism*. Moscow: Progress.

Leont'ev, A. N. (1978). *Activity, Consciousness and Personality* (M. J. Hall, Trans.). Englewood Cliffs, N. J.: Prentice-Hall.

Leont'ev, A. N. (1981). *Problems of the Development of the Mind*. Moscow: Progress.

Nardi, B. M. (1996a). Some Reflections on the Application of Activity Theory. In B. M. Nardi (Ed.), *Context and Consciousness: Activity Theory and Human-Computer Interaction* (pp. 235-246). Cambridge, Ma.: MIT Press.

Nardi, B. M. (Ed.). (1996b). *Context and Consciousness: Activity Theory and Human-Computer Interaction*. Cambridge, Ma.: MIT Press.

Sanderson, P. M., & Fisher, C. (1994). Exploratory sequential data analysis: Foundations. *Human-Computer Interaction, 9*(3), 251-317.

Vygotsky, L. S. (1978). *Mind in Society*. Cambridge, Ma.: Harvard University Press.

Vygotsky, L. S. (1986). *Thought and Language* (A. Kozulin, Trans. Revised Ed ed.). Cambridge, Ma.: The MIT Press.

Wertsch, J. V. (Ed.). (1981). *The Concept of Activity in Soviet Psychology*. Armonk, NY: M.E. Sharpe.

Wertsch , J. V. (Ed.). (1985). *Culture, communication and cognition: Vygotskian perspectives*. Cambridge: Cambridge University Press.

Wertsch , J. V., del Rio, P., & Alvarez, A. (Eds.). (1995). *Sociocultural Studies of Mind*. Cambridge: Cambridge University Press.

Zinchenko, V. P., & Gordon, V. M. (1981). Methodological Problems in the Psychological Analysis of Activity. In J. V. Wertsch (Ed.), *The Concept of Activity in Soviet Psychology*. Armonk, NY: M.E. Sharpe.