Towards a user-centred method for studying CVLEs

*Daphne Economou, William L. Mitchell and Tom Boyle**

Department of Computing and Mathematics,
Manchester Metropolitan University,
John Dalton Building, Chester Street,
Manchester, M1 5GD
Tel.: ++44 (0)161-247-1492,
Fax: ++44 (0)161-247-1483
E-mail: {D.Economou, B.Mitchell}@doc.mmu.ac.uk
URL: http://www.doc.mmu.ac.uk/RESEARCH/virtual-museum/Kahun/

*School of Informatics and Multimedia Technology,
University of North London,
Faculty of Science, Computing and Engineering,
166-220 Holloway Road,
London, N7 8DB
E-mail: T.Boyle@unl.ac.uk
Tel.: ++44 (0)171 973 4810,
Fax: ++44 (0)171 753 7009

Abstract. This paper presents a method for studying multiuser systems in an educational context. The method has been developed as part of the Senet project, which is investigating the use of virtual actors in Collaborative Virtual Learning Environments (CVLEs). A groupware prototype has been studied in order to identify requirements and design factors for the Collaborative Virtual Environments (CVEs). Data collection occurs by videotaping users, capture of text typed in chat boxes, and note taking. The analysis part of the method is partly based on Interaction Analysis and results in a mixture of quantitative and qualitative findings. It deals with transforming vast amount of rich qualitative data in a quantitative form, that can be used to draw design principles for constructing CVEs for learning. The method is described by second phase of work in the Senet project. The paper also shows how the method is being extended to provide a more summative evaluation of the CVE to be developed.

Keywords: requirements gathering, evaluation, Interaction Analysis, museum education, virtual environments, groupware

1. INTRODUCTION

This paper describes a method for studying Collaborative Virtual Learning Environments (CVLEs). The method has been developed as part of the Senet project (see section 2) which studies the role of virtual actors in Collaborative Virtual Environments (CVEs). It aims to understand the way that CVEs for learning are formed. It identifies the design factors involved in the construction of CVEs for learning and the respective properties that these are formed. (Stephanidis *et al.*, 1999) elucidate such an approach for studying "new virtualities" (multiuser virtual communities). A phased approach is adopted where robust prototypes are constructed and studied in order to determine requirements and design factors for CVEs.

The second phase of the Senet project involved the construction of a prototype based on groupware technology. The method described in this paper was developed to study the prototype in this second phase. Methods such as Conversational Analysis and Discourse

Analysis provide too fine-grained analysis for the purposes of requirements gathering. For this reason a 7-step method was proposed which will meet these needs as well as leading to a set of design factors (see section 3). Each step of the method is shown with examples from the second phase studies (see section 4).

The method is intended to be used for more summative evaluations of a full CVE in the third phase of the Senet project (see section 5). By extending it the method will be applicable in a wider range of situations.

2. BACKGROUND TO THE STUDY

2.1. Aims of This Research

The aim of this research is to investigate the role of virtual actors in CVEs for learning. It develops a framework of design factors for the use of virtual actors. These factors will address issues of pedagogy and interaction (e.g. how can actors be used to structure a learning session in virtual environments) rather than underlying technological concerns (e.g. the suitability of particular computer graphics rendering techniques). The adopted approach here identifies the design factors influencing the effectiveness of an educational CVE. For this purpose it is first necessary to study the interactivity and social communication arising in such collaborative learning situations.

2.2. Problems

There are three main problems in the development of design guidelines for the use of virtual actors in CVEs:

- the current prototypical nature of many of the applications developed

  Current studies have been with users with ready access to the technology. However, the recognition of the situated nature of the processes that arise in collaborative learning is necessary. (Roussos *et al.* 1999) calls for more exploratory work which involves building novel learning applications and carrying out informal evaluations of them.
- the vast amount of factors that are involved in the construction of CVEs

  Kaur (1997) has identified 46 design properties to be considered when designing VEs for usability. The number of factors increases dramatically when considering communication and collaboration issues. It is then difficult to isolate which design decisions are responsible for the overall effectiveness of the environment and identify the inter-play between various factors (e.g. the effects that usability issues have on pedagogic issues).
- the current immaturity of the technology underlying CVEs for learning

  The full potential of the CVLE medium cannot be exploited (e.g. the range of interaction and image quality is often limited).

2.3. Towards a methodology for studying CVEs for learning

The overall research methodology followed in this research has been presented elsewhere (Economou *et al.*, to appear). The two following sections briefly summarise some important points concerning the methodology addressing the above problems.

2.3.1. Real-world application

To address the first problem (see section 2.2) and determine the requirements of a CVE it is necessary to study a “real world” situation. Only in such a situation seemingly trivial problems arise that in reality may determine the success or failure of a system.

In order to study an authentic learning activity the research was based around the work of Manchester Museum's Education Service (Mitchell, 1999). This service caters for school visits to the museum aimed at Key Stage Level 2 (9-11 years old). It provides access to a wide range of Museum artefacts relevant to subjects in the National Curriculum for education. One particular strength of the Museum with a major part in the Education Service's teaching, is its collection of every day life ancient Egyptian artefacts from town of Kahun. The artefact chosen as basis of the learning activity in this research is Senet - a board game for two players. Players take turns to throw a die. The object of the game is to "bear off" your 10 pieces first. Through the activity and a collaborative process the children get familiar with the artefact and learn by using it how it was played (see Figure 2(b)).

A CVE based on Senet provides a good testbed for various CVLE properties. Participants can manipulate objects in the environment such as board, die and pieces. It allows participants to operate as individuals, in pairs or as larger groups. In terms of collaboration it allows co-operation (to learn the game) as well as competition (to win the game). The game situation also allows a range of teaching styles from traditional instructional methods (e.g. explaining the rules in advance) to constructivist methods (learning by playing).

The game supports the needs of experimentation in various ways. It can be implemented in either 2D or 3D form. This allows the construction and study of different forms of prototypes. It is a fairly well structured task (the players have to follow certain steps to learn the rules and play the game). The length of the required time matches well the length of time the children could take part in a task before becoming restless (30-45 minutes). Assessment of the player’s knowledge can occur in a fairly unobtrusive manner (e.g. by observing if they are following the rules).

2.3.2. A phased approach

To address the second problem (see section 2.2) and simplify the design factors’ study a phased approach was adopted (Economou *et al.*, 1998). To deal with the third problem (see section 2.2) more mature technologies were used to construct robust prototypes. This allows essential features of the situation (interaction and communication) to be studied with real users in order to identify the types of interactivity and social communication that need to be supported in CVLEs.

In the first phase, a prototype environment was developed with the form of a single display groupware (Stewart *et al.*, 1999). Participants sit next to each other and view the environment through a single, shared display. The interactions between them were external to the computer (see Figure 1(a)). The prototype was constructed using 2D-multimedia technology (Macromedia Director). This simplified issues of the navigation of the environment and the ways in which objects could be manipulated. The purpose of this phase was primarily exploratory in nature. It was meant to gather information about the "real world" and identify usability issues of the prototype in order to inform the design of environments developed in subsequent phases.

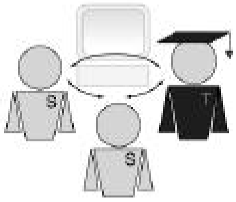

Figure 1(a). Interactions external to the environment

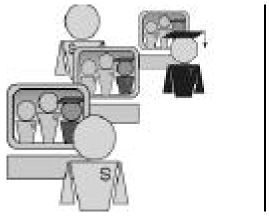

Figure 1(b). Interactions internal to the environment

The prototypes developed in the second phase took the form of conventional groupware systems. Participants were remotely located so interactions between them were internal to the computer (see Figure 1(b)). The prototypes were developed using 2D multimedia tools coupled with groupware technology (NetMeeting). NetMeeting supports sharing of documents and applications and synchronous communication via text or audio. The purpose of this phase was to explore the interaction being internal to the environment and the effects on the behaviour of participants. The prototypes also introduced the concept of population to the environment. One prototype was semi-populated (the child could see a virtual actor representing the expert) and the other two prototypes were fully populated (the children could also see their own virtual actor).

In the third phase, a prototype is being constructed using CVE technology. This will be 3D, fully populated with the interactions internal to the computer. The prototype design is according to the design factors identified in previous phases. The studies in the first two phases have exploratory nature. The studies outcome was a rich set of qualitative information used to form a set of design factors that informs the use of virtual actors in CVEs. The evolving design factors then guide the development of the CVE prototype in the third phase. The work in the third phase involves evaluation of a more summative nature.

The phased approach provides several benefits: (i) provides means of managing the complexity of factors by dealing with a manageable set of factors in each phase (e.g. 2D/3D and population) and (ii) it allows the results of each phase to inform subsequent phases. Thus it allows requirements to be progressively identified.

3. A METHOD FOR STUDYING CVLES

3.1. The purpose of the study and the prototypes studied

This paper focuses on the method used to study internal to the environment interaction and communication between participants.

A set of groupware prototypes were developed *ad hoc*. Each prototype contained three main elements: game artefacts, user representations and communication tools. Participants had to communicate with each other via the communication tools. These consisted of chat boxes (for typed communication), a hand (for pointing things in the game) and a white board (for other communication e.g. drawing). The communication tools were designed to emulate the type of tools typical for CVEs.

Three different prototypes were developed and observed:

1. 2-D semi-populated, dialogue external to the game environment (see Figure 2(a))
   The first prototype consisted of a child and an expert (played by a researcher) located in different rooms. The child could see virtual actor representing the expert but was not represented by own virtual actor. Participants communicated by typing text in a NetMeeting chat window which was external to the game environment window.
2. 2-D fully-populated, dialogue internal to the game environment
   Each participant was represented by own virtual actor (a palette allows selection). In addition, participants communicated via chat boxes which are internal to the game environment. This represents an environment in which a "speech bubble" is associated with each actor.
3. 2-D fully-populated, dialogue internal to the game environment (see Figure 2(b))
   This differs from the second prototype as second child is present in the environment.

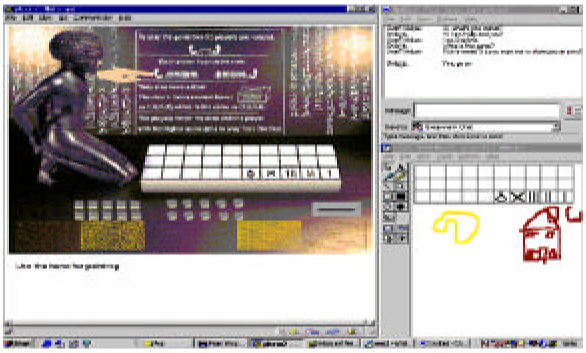

Figure 2(a). 2-D semi-populated, dialogue external to the game environment

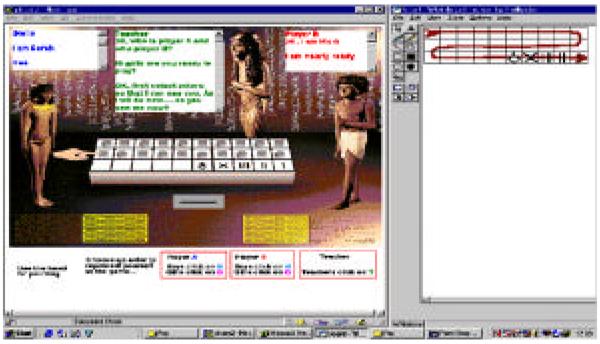

Figure 2(b). 2-D fully-populated, dialogue internal to the game environment

The method adopted for this study is primarily for requirements gathering. Thus the study is not meant to be as exhaustive and fine-grained as more sociological or linguistic studies (see next section). It is also important to provide some degree of objectivity in terms of support for the proposed design factors.

3.2. Different ways of studying social interaction

Methods studying social interaction were developed in several research areas. (Stephanidis *et al.*, 1999) argues that facilitating the design of “new virtualities” is drawn upon the accumulated knowledge and results of established theoretical stands within social sciences and /or psychology. Ethnomethodology and conversation analysis made profound contributions to understanding the practices and reasoning through which individuals produce social actions and activities and recognise the activities of others (Atkinson & Heritage, 1984; Boden & Zimmerman, 1991; Silverman, 1997). Other methods derive from linguistics. For example Discourse Analysis, as practised by the Birmingham group, evolved as means of studying structure in classroom discourse (Coulthard *et al.*, 1981). Understanding of the body in social action and interaction is mainly done in psychology regarding ‘non-verbal behaviour’ (Cassell, to appear).

The use of audiovisual equipment provides a flexible source of rich data for analysis as it captures actions and activities accomplished through the body and physical artefacts, as well as talk in face to face interaction. (Heath, 1997) argues that video coupled with a suitable methodology framework such as ethnomethodology and talk analysis can also provide means for studying the visual and vocal aspects of human conduct. Although sociology has been slow in exploiting the opportunities of the use of video, traditional social anthropology such as ‘interaction analysis’ uses recorded data for naturalistic activities (Scheflen, 1964).

One method for which video technology is essential is ‘Interaction Analysis’ (capitalised) (Jordan & Henderson, 1995). This method has its roots in the social sciences and sees knowledge and action as fundamentally social in origin, organisation and use. It studies human activities, such as talk, non-verbal interaction and the use of artefacts and technologies. It is primarily defined by its ‘analytic foci’ or ways into a video-tape. Such foci include: structure of events, temporal organisation of activity, turn-taking, trouble and repair, spatial organisation of activity. Important to Interaction Analysis is the data analysis by a group of analysts. The integrated features of the method made it particularly suitable as a base for the Senet project.

3.3. The need for a grid-based method

The primary purpose of the method to be adopted is as a form of requirements gathering to enable the identification of design factors. Methods such as Conversation Analysis and Discourse Analysis are somewhat narrowly focused on issues surrounding the dialogue itself. An ethnographic approach, particularly when coupled with video will result in a vast amount of rich qualitative data gathered over a long period of time. This is not only unmanageable but the moment-to-moment detailed analysis required is a very long process. The information is interrelated and it is difficult to be separated and rationalised. (Viller, 1999) argues that is difficult to draw design principles and other abstract lessons from a technique that is concerned with detail of a particular situation. Thus, it is difficult to make generalisations about design factors related to virtual actors. The analysis itself can also appear very subjective as it is based on the experience of individual analysts.

The group based analysis advocated by Interaction Analysis goes some way to countering subjectivity of analysis. However, group based analysis is not always possible (as in the case of the Senet project) because of resource limitations.

The proposed solution that addresses these problems is the creation of an analytic grid that can be used to generate numerical values from the qualitative data (see section 4.4, 4.5). For example if the factor to be studied is physical activities that the virtual actors should be eligible of to improve communication and interaction issues, then the quantitative information subtracted out of the qualitative data should indicate in which circumstances, and for what purpose certain physical activities have been used. In this form the data is much more manageable and can be linked forward more reliably to the design factors that are developed.
The analytic foci and orientation adopted in the method used to study the Senet project, outlined next, is based and adds on the Interaction Analysis foci.

3.4. The outline of the method adopted for studying the Senet project

The adopted method studying the Senet project consists of 7 main steps:

- data collection
- transcription
- chunking of the transcription
- creation of a grid

- application of the grid
- analysis at the session level
- derivation of design factors

The section below outlines the use of the method in the second phase of the Senet project.

4. THE METHOD APPLIED

4.1. Data collection

The second phase took place at a local high school over three school days. The subjects were eleven-year old children (year 7). Twenty two children (11 pairs) participated in the studies. Two rooms were used (see Figure 3). One contained a researcher playing the role of the 'expert' and the second contained one or two children working on individual computers accompanied by a second researcher (the helper). In the first and second studies only one child used the environment. The second child adopted the role of an observer.

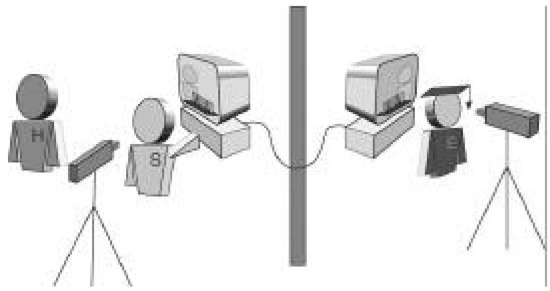

Figure 3. The physical set up of the study
H=helper providing technical support, S=the child, E=the researcher playing the role of the expert

Questionnaires filled out by the children before the session obtained background information about the children. Each session lasted approximately 45 minutes. Basic instructions about the system were given the children at the start of the session and were instructed to ask the expert for support. The expert and the child were video taped separately. The text the participants exchanged during a session was written into a file. Each session was followed up by an interview of the children about their experiences, that lasted approximately ten minutes and was tape recorded.

Video technology provided a permanent record of events where language, gestures and interactions between people and between people and artefacts were captured. The recorded data can be used for both qualitative and quantitative analysis. It can be played back repeatedly and shared with colleagues. It can be seen as more objective because it offers scope for re-interpretation and avoids the idiosyncrasy of individually recorded field notes. This means that the validity of any findings can be easily revised and verified. However, although video provides rich and valuable data it can be overwhelming, particularly as two video tapes were produced for each session.

4.2. Transcription

The transcription step involves creating one account of a session by combining the data collected about the interactions taking place internally and externally to the prototype. The starting point of the recording is the text typed in the chat boxes. This generates the internal to the prototype dialogue, including the time the text typed and the participant responsible.

The two videotapes are synchronised, based on events noticeable on both videotapes (e.g. sound, a change on the screen) and a description of all actions is created. This includes both internal information (e.g. moving pieces, using the hand or the drawing tools), external information (e.g. dialogues between people in the real world; movements and gestures). This transcription provides a more manageable way handling the data (see Figure 4). For the purpose of this paper the same example session is used.

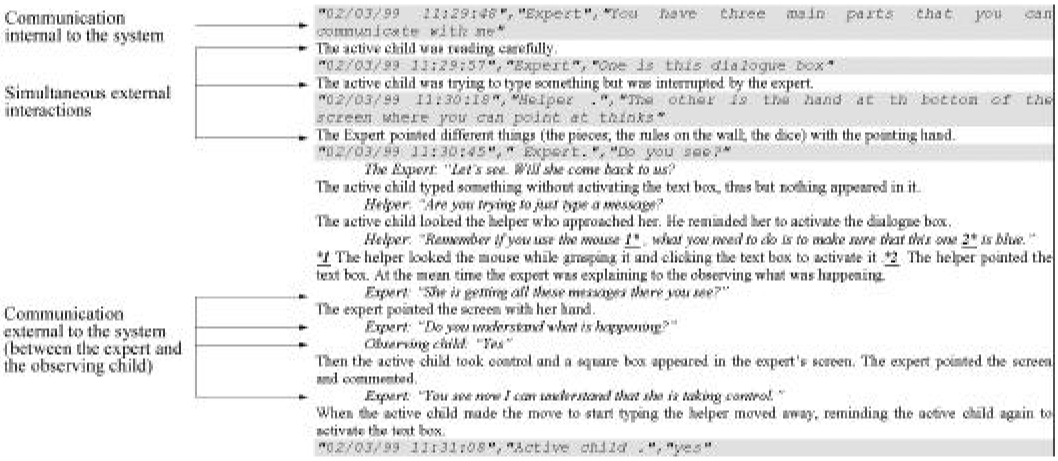

Figure 4. An illustration of a part of the example session transcription

4.3. Chunking of the transcription

Dividing the transcription into a series of ethnographic chunks provides a more manageable set of units for analysis. According to Interaction Analysis, one such chunk is the *event*, which is a stretch of interaction that coheres in some manner that is meaningful to the participants. Events can be named and constitute recognisable, culturally significant signs in social communication. (Jordan & Henderson, 1995) refer to tutoring sessions, bedtime stories, as examples of recognisable events. Events in turn can be sub-divided into a set of *segments* (e.g. in a meal event segments such as 'setting up the table' or 'serving the coffee' can be identified. To identify ethnographic chunks it is necessary to draw on cultural knowledge or local experts. For this particular study 4 ethnographic chunks were identified:

- session
- stage
- segment
- turn

The *session* is equivalent to a game playing event. A session consists of a series of *stages*. The structure of a session depends on the teaching strategy adopted based on how the situation unfolds. For example, the session to be analysed below consisted of 6 stages (see Table 1). In other sessions, stage 5 was replaced by 2 separate stages, like the expert presenting the rules followed by the game being played.

1.	the helper gives some introductory information about the following activities and explains to the child how to communicate with the expert using the chat window that the system provides
2.	the expert takes over and verifies that the child knows how to communicate with him/her using the chat window
3.	the expert explains the interface to the child using the provided communication tools
4.	the expert explains the first part of the rules about setting up the board
5.	the expert explains the second part of the rules about playing the game by playing with the child
6.	the session concludes

Table 1. Stage structure of the example session

Stages are sub-divided into segments. *Segments* are differentiated by a change in topic. For example, in the first stage (helper gives some introductory information) two segments can be identified (see Figure 5). In the first segment the helper explains to the child how communication takes place with the expert. In the second segment the helper provides some additional information to cover the delay while the expert replies.

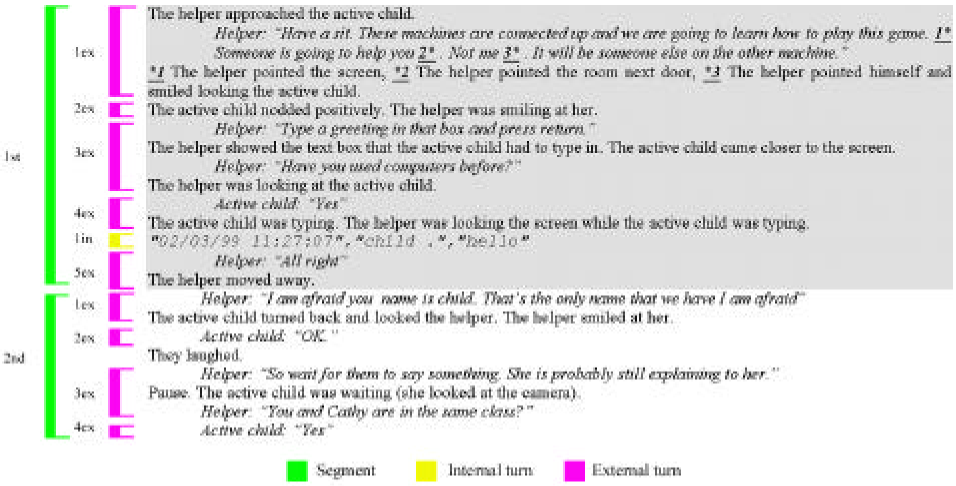

Figure 5. Two segments of the 1^{st} stage of the example session

The last used ethnographic chunk unit was the *turn*. The boundary of a turn is marked by the other participant taking control. There are two kind of turns: internal (a turn using one of the prototypes tools e.g. typing in a chat box) and external (between participants external to the prototype). In the above example of two segments (see Figure 6), one internal and nine external turns can be identified.

4.4. Creation of a grid

4.4.1. Forming a grid

The moment-to-moment detailed analysis of each ethnographic chunk leads to a vast amount of rich qualitative data in an unmanageable form that generalisations about design factors are difficult. To address this problem the development of a 'grid' is proposed, that aims to generate countable values out of the rich qualitative data. The grid is formed of analytic categories based on the studied factors. The extracted quantitative data provides a more concise set of findings that can then be used to derive design factors.

To identify the features to appear in a grid it is necessary to have a clear idea about the studied factors. This might require a hypothesis at the start of the study. The exploratory nature of this research makes this approach impossible.

The first research phase provide a solution. The outcome was a framework of design factors based on studies carried out with prototypes as well as an analysis of existing CVEs (Benford *et al.*, 1995; Capin *et al.*, 1996). The framework included factors such as:

- ***appearance*** (e.g. user's representation)
- ***awareness***, what the actor can perceive about the VE and situation
- ***object manipulation***, ways in which the virtual actors can use objects (e.g. moving a piece in the senet game)
- ***communication content***, the content of the communicative exchanges
- ***communication modes***, the ways in which the virtual actors communication can be presented (e.g. as text or as speech)
- ***turn-taking***, this may be by taking turns at communicating (e.g. interrupting) or object manipulation
- ***role in situation***, this will depend on the particular pedagogical style being followed

One problem was that the data from the first phase relates to external to the prototype interaction, whereas the second phase with internal interaction. If the framework was followed strictly some important factors could be ignored. To overcome this a preliminary analysis of selective second phase transcriptions was carried out. These preliminary results were combined with the design factors framework to create the analytic categories of the grid.

4.4.2. The grid used in the study

The analytic categories identified were:

- physical activity
- communication activity
- turn taking
- external intervention
- pedagogy

The columns in the grid differ depending on the particular analytic category (see Figure 6).

However there are some columns common to all the grids:

- chunk index
- (e.g. the chunk that is first stage, second segment, fifth internal turn is coded as 1b5in)
- location of the action
- This indicates the relative use of communication tools (internal) vs. external communication prototype)
- description of the chunk
- participants involved

Physical Activity	type of physical activity											
Chunk index	Head movement		Facial expression	Position of the body	Movement of the rest of the body (excluding head)				Location of the action		Description of the chunk	Participants
	looking	nodding	happy, sad, worried, uncertain		walking	pointing	moving object	gesturing	internal interaction	external interaction		

Communicative Activity	type of communicative activity													
Chunk index	Text	Pointing			Drawing	Speech	Body language				Location of the action		Description of the chunk	Participants
		external pointing	cursor	hand tool			head movements	facial expressions	gesturing	whispering	internal interaction	external interaction		

Turn taking											
Chunk index	Turn boundaries			Force			Interruption mechanisms	Location of the action		Description of the chunk	Participants
	textual	audio	visual	command	question	statement		internal interaction	external interaction		

External intervention												
Chunk index	How participant roles change			Prompt for the intervention	Reason for the intervention	Intervener's action		Location of the action		Description of the chunk	Participants
	helper	expert	active child			to recover	to prevent	internal interaction	external interaction		

Pedagogy									
Chunk index	Participant adopting teacher's role	Topic	Pedagogic style	Change in style	Reason for style	Location of the action		Description of the chunk	Participants
						internal interaction	external interaction		

Figure 6. The grid

4.5. Application of the grid

To illustrate the use of the grid, the same example session transcription is analysed (Figures 4, 5 and Table 1 are referred to the same example session transcription as well) in turn for each of the analytic categories. This example involved a helper and an active child in one room communicating with the expert remotely. An observer child accompanied the expert. The first prototype was used (see Figure 2(a)).

The analytic category of **physical activity** concerns the physical movements of the users. There are four main types of physical activity:

- head movement (e.g. where the participants are looking, nodding)
- facial expression
- position of the body
- movement of the rest of the body excluding the head (e.g. walking, pointing, moving objects and gesturing)

This analytic category is examined at the turn level.

The grid can be applied to the example session transcription to examine position of the body. There were six cases that the helper stood close to the child to provide support. Three times the active child and once the observing child, moved closer to the screen when the expert asked them a question to assess their understanding.

In terms of other body movements pointing occurred 34 times. Of these occurrences 29 were used to complement spoken or written dialogue and 5 on their own (as a means of reply). Of the 34 examples of pointing, 13 involved pointing internal to the prototype, and 21 external pointing. For external pointing, the helper used pointing to show the active child where to type in the chat window, while the used expert pointing to show the observer child where the active child's dialogue would appear. For internal pointing, the expert indicated the tools that could be used by the active child and in 4 cases underlined the parts of the rules on the wall that the child needed to read.

Thirty six times objects were moved around the screen, including moving pieces on or off the board, rolling the dice and scrolling the rules on the wall. Ten times moving an object was used as a reply. Twice the expert moved pieces to correct the active child's move. No sufficient data is available to make observations related to head movement and facial expression because the cameras were set up to capture the screen only.

The analytic category of **communication activity** considers the mode of communication:

- text (via the chat boxes)
- pointing (via the cursor or hand tool)
- drawing
- speech (between participants external to the prototype)
- body language (head movements, facial expressions, gestures, whispering)

This analytic category is examined at the turn level.

Application of the grid to the example session transcription shows 42 turns using text mode. Thirty four turns were pointing. Sixty five were speech turns, 23 of these were between the helper and active child, and 42 between the expert and the observing child.

The description column of the grid includes the content of the turns. For example, between the helper and the active child these included: introduction to the session; getting information about the active child's background; responding to questions; reminders; providing extra help; confirmation and feedback about actions. Between the expert and the active child the following can be identified: the expert giving commands; assessing; explaining the interface and the rules of the game (in many cases accompanied by pointing); reminders; demonstrations; feedback, correction of moves and suggestion of alternative moves.

In **turn taking**, the issues to be studied are:

- turn boundaries (marked by textual, audio, or visual markers)
- force (command, question or statement)
- interruption mechanisms

This is examined at the turn level.

Using the example session transcription, 35 turn boundaries were marked textually, 27 by sound, and 37 visually. Textual boundaries occur in the chat window. Sound boundaries occur in the speech external to the prototype. Of the 37 visual boundaries 8 occurred externally (participants can see each other) and 29 internally (marked by participants moving objects). Considering the textual and sound turns, 9 were commands (5 textual and 2 speech), 28 questions (11 textual and 17 speech) and 25 statements (13 textual and 12 speech).

The nature of the technology means that participants have equal access to the pointer. There is thus no explicit interruption mechanism. The only scope for interruption was for the helper to interrupt the active child. One notable exception in the example transcription occurred when the observer child interrupted the expert. This was to point out a mistake that the active child made but that the expert missed.

The analytic category of **external intervention** concerns complete break down in a session. It addresses cases where intervention (usually by the helper) is required in the real world, external to the prototype. It is concerned with investigating why the expert's support wasn't sufficient and additional external help was required. It studies:

- how participant roles change
- prompt for the intervention (what signals led to the intervention)
- reason for the intervention
- intervener's action

(e.g. to recover from breakdown or prevent serious breakdown)

This is examined at the session or segment level.

In the example session transcription external intervention occurred seven times. Six times the helper adopted the role of the expert. In one case the active child requested the helper's support. In the rest of the cases the helper was prompted through awareness of external conditions. Once the observing child prompted through awareness of external conditions as well as internal once and adopted the role of the expert to remind the expert to correct the active child. In the example, external intervention always occurred due to a combination of technical problems. This was due to the lag in the system response. This made it difficult to distinguish genuine lag from an extended turn. Another issue was lack of knowledge about the interface to the communication tools e.g. knowing to click in a chat box to make it active, or moving the mouse and thus stopping the other participant from gaining control of the shared cursor. The helper's actions were always to help the child recover from the breakdown.

The last analytic category of the grid is **pedagogy**. It studies:

- participant adopting the role of the teacher
- topic
- pedagogical style used
- change in style
- reason for the style being used

A pedagogic style can only be followed across an entire segment or even stage. This is due to the fact that a certain pedagogical tactic cannot be identified by a single turn but throughout several turns. This analytic category is thus applied at the stage or session level.

In the example session transcription, the helper and the expert followed almost the same pedagogical style in the first three stages. This involved explaining something to the child and asking the child to try it (e.g. 'this is where you need to type', 'type something', or 'we use the pointing hand to show things to each other', 'show me the dice'). In the fourth stage the style changed as the expert had to explain bigger sections of information to the child (e.g. the rules of the game). The expert in this case used a system tool to show and demonstrate things and kept long turns. In the fifth stage, the style changed as the child was learning the rules by playing. The expert took the role of the knowledgeable partner who explained how the game was played and at the same time played the game with the child.

4.6. Analysis at the session level

The use of the grid provides a set of numerical values at the turn or segment levels. This step is about investigating what these figures reveal at the session level.

The analyst needs to study the flow of the segments in the whole session based on the timeline that the grid provides. The numerical values and descriptions from the grid are used to provide a more concise description of the session than that provided by the transcription. This allows examination of the overall structuring of the session. The structure of the example session has been described previously (see Figure 5).

The analyst also needs to examine how and why the participants' behaviour changes over the whole session. For example, turn taking can be examined across the whole session. The expert and the active child take turns using a protocol in which the expert asks a question or gives a command to the child and the child answers using the chat box. When the expert has to explain more information uses extended turns. The child after several unsuccessful attempts to interrupt the expert stops attempting to take control or typing a message. From now on the child replies by pointing or moving objects only when is asked to do so.

Also different analytic categories for the same session needs to compare. The grid chunk indexing and descriptions help to facilitate this. For example, comparing the physical activity and communication activity categories for the example reveals cases in which pointing or moving objects is used to complement chat window dialogue.

The final stage is to compare the findings for sessions against each other in order to identify patterns throughout all the sessions.

There is a possibility that this step might identify that there are problems with the grid itself, in terms of the analytic categories that the grid deals with. It might identify other analytic categories need to be added to the grid and investigated further.

4.7. Derivation of design factors

In this step the results of the analysis are used to develop design factors.

Using the results from the example analysis, the physical and communication activities grids indicate that pointing (both internal and external) is crucial as it acts as a complement to communicated dialogue as well as a means of communication itself. Body positioning was also important as participants moved closer to each other to speak or provide extra help. This indicates that in the third phase of the Senet project virtual actors should be able to:

- walk in the CVE, indicating the participant's position in the environment
- select an object (if the virtual actor is "within arm's reach" of the object then the virtual actor positions its hand correctly to "touch" it) or to pick a distant object by pointing (in which case a "laser pointer" from the hand indicates selection).

The third phase of the project where DEVA CVE (Pettifer, 1999) technology is used addresses these issues as the technology supports these features.

The communication activities and the turn taking analytic categories, plus the flow of the events of the grid timeline, reveal that the ratio of the number of turns to amount of topics covered in the external communication was far greater than the internal communication. Based on the performance of the active and the observing child it seems that the observing child had a better understanding of the rules of the game despite her passive role. These are two of the issues that indicate the effectiveness of speech. In the post-session discussion with the children it was emphasised that speech would provide a more direct way of communication. Thus, a design factor for the development of the virtual actors is to support speech.

However, text is very important for the specific application, because a large set of rules has to be presented to the players. Additionally, skills such as reading and writing are of a significant importance in a learning environment and are strongly supported in the National Curriculum. One issue not clear in the example transcription but very clear in other sessions is that use of the chat box as "transcript" of the dialogue, used both by the expert and the children (e.g. scrolling back over an explanation) to kept track of previous actions and dialogues. All these issues show that the virtual actors should be provided with a combination of different means of communication to be used according to the activity.

Development of tools that let the virtual actors keep track of the history of the users' activities is also needed. This is a direction taken for the third phase of the project.

The study of turn taking, external intervention and pedagogy shows the importance for the expert to be aware of and to be able to control the activity. The virtual actors can provide other clues about the current focus of the activity of the participant (e.g. if they are speaking or moving an object) as well as an intention to interact (e.g. actors can approach each other to speak). The representation of the virtual actors should address these problems. For example, the participants could be provided with tools to interrupt and ask a question (e.g. a question mark could appear over their head, or a message could appear in all participants' chat boxes). These are just brief examples of how some of the analyses based on the example transcription could lead to design factors for the use of virtual actors.

5. CONCLUSION

The proposed method provides a means of managing a large amount of disparate data (2 video tapes, audio, notes, text files). The grid provides a means of obtaining a more concise and objective measure of the moment-to-moment details occurring in a session. The chunking, indexing and use of the timeline provides means of looking at the overall structure and flows within a session.

There are some drawbacks with the method. Firstly, it is not as exhaustive or generates as much rich, qualitative informative as, for example, ethnographic techniques. However, the primary purpose of the method is to gather requirements within certain restrictions of time and resources. Secondly, the grid is more suitable for analysing at the turn level rather than for broader issues such as pedagogy, where a whole segment needs to be analysed to reflect the use of different pedagogic tactics in a stage. Further development is needed to the method to address this issue. Thirdly, the factors to be studied must be reasonably clear in order to derive analytic categories for the grid. This means that in the early stages of research some exploratory studies are needed (as was the case in the first phase studies in the Senet project).

The next phase of work in the Senet project is to develop a full CVE prototype according to the design factors generated so far. In this phase the method will then be used mainly in an evaluative fashion. Saying that, the situation to be evaluated will be different from the situation examined so far. It is therefore important not to ignore new issues that may arise.

The method developed shares many of the analytic foci that defines Interaction Analysis. It builds on this, by means of the grid to provide a more efficient and rigorous requirements gathering technique. The next phase of work will show how the method can be extended to form an evaluation method for CVEs.

ACKNOWLEDGEMENTS

Thanks to the State Scholarships Foundation of Greece (S.S.F.) for funding Daphne Economou's Ph.D.; the Advanced Interfaces Group at the University of Manchester Department of Computer Science for help with the third phase and Knutsford High School for their co-operation.

REFERENCES

Work from the Senet project can be found at: www.doc.mmu.ac.uk/RESEARCH/virtual-museum

Atkinson, J.M. & Heritage, J. (1984). *Structures of Social Actions: Studies in Conversation Analysis*, Cambridge: Cambridge University Press.

Benford, S., Bowers, J., Fahèn, L., Greenhalgh, C. & Snowdon, D. (1995). User Embodiment in Collaborative Virtual Environments, *Proceedings of ACM CHI'95 Conference on Human Factors in Computing Systems,* ACM Press, pp.242-249.

Boden, D. and Zimmerman, D.H. (1991). *Talk and Social Structure: Studies in Ethnomethodology and Conversation Analysis*, Cambridge: Polity.

Capin, T.K., Pandizc, I.S., Chauvineau, E., Thalmann, N.M. & Thalmann, D. (1996). *D4.1, Report on Modeling and Animation of Virtual Humans, COVEN – Collaborative virtual environments*, CEC Deliverable Number: AC040-GEN-MIR-DS-P-041.b0.

Cassell, J. (to appear). Embodied Conversation: Integrating Face and Gesture into Automatic, Spoken Dialogue Systems, Luperfoy (ed.), *Spoken Dialogue Systems*, Cambridge, MA:MIT Press.

Coulthard, M., Montgomery, M., and Brazil, D. (1981). Developing a Description of Spoken Discourse, *in Studies in Discourse Analysis*, M. Coulthard and M. Montgomery (eds.), Routledge & Kegan Paul, London, England, 1981, pp.1-50.

Economou, D., Mitchell, W.L., & Boyle, T. (1998). A Phased Approach to Developing a Set of Requirements for the Use of Virtual Actors in Shared Virtual Learning Environments, *Proceedings of ED-MEDIA & ED-TELECOM 98: 10th World Conference on Educational Multimedia and Hypermedia & World Conference on Educational Telecommunications*, Association for the Advancement of Computing in Education (AACE), pp.1620-1621.

Economou, D., Mitchell, W. L., Boyle, T. (to appear). Requirements Elicitation for Virtual Actors in Collaborative Learning Environments, *Computers & Education*, R. S. Heller & J. Underwood (eds.), Elsevier Science Ltd., Oxford.

Heath, C. (1997). The Analysis of Activities in Face to Face Interaction Using Video, *Qualitative Research: Theory, Method and Practice*, Silverman, D. (ed.), London, Sage, pp.183-200.

Jordan, B. & Henderson, A. (1995). Interaction Analysis,: Foundations and Practice, *The Journal of Learning Sciences*, Lawrence Erlbaum Associates, Inc., Vol.4, No.1, pp.39-103.

Kaur, K. (1998). *Designing virtual environments for usability*, PhD thesis, Centre for HCI Design, City University, London.

Mitchell, W.L. (1999). Moving the museum onto the Internet: The use of virtual environments in education about ancient Egypt, J.A. Vince and R.A. Earnshaw (eds.), *Virtual Worlds on the Internet*, IEEE Computer Society Press, pp.263-278.

Pettifer, S. (1999). *An operating environment for large scale virtual reality*, PhD Thesis, The University of Manchester.

Roussos, M., Johnson, A., Moher, T., Leigh, J., Vasilakis, C., & Barnes, C. (1999). Learning and Building Together in an Immersive Virtual World, *Presence*, 8(3), pp.247-263.

Scheflen, A.E. (1964) The significance of posture in communication systems, *Psychiatry*, 27, pp.316-31.

Silverman, D. (1997). *Qualitative research:theory, method and practice*. Silverman, D. (ed.), London, Sage.

Stephanidis, C. (Ed.), Salvendy, G., Akoumianakis, D., Arnold, A., Bevan, N., Dardailler, D., Emiliani, P.L., Iakovidis, I., Jenkins, P., Karshmer, A., Korn, P., Marcus, A., Murphy, H., Oppermann, C., Stary, C., Tamura, H., Tscheligi, M., Ueda, H., Weber, G., & Ziegler, J. (1999). Towards an Information Society for All: HCI challenges and R&D recommendations. *International Journal of Human-Computer Interaction*, 11 (1), pp.1-28.

Stewart, J., Bederson, B.B., & Druin, A. (1999). Single Display Groupware: A Model for Co-operative Collaboration, *Proceedings of ACM CHI'99*, Conference on Human Factors in Computing Systems, ACM Press.

Viller, S. and Sommerville, I. (1999). Social analysis in the requirements engineering process: from ethnography to method, *Proceedings of International Symposium on Requirements Engineering*, Limerick 1999, IEEE Computer Soc. Press, pp. 6-13.