Activity Theory: Another Perspective on Task Analysis

Phil Turner and Tom McEwan

Social Computing Research Group,

School of Computing, Napier University

10 Colinton Road, Edinburgh, EH10 5DT

p.turner@napier.ac.uk t.mcewan@napier.ac.uk

In this chapter we introduce activity theory, describing its origins and its principal researchers and thinkers. We stress its descriptive power and its usefulness in defining uniquely a unit of analysis for work. We contrast its development with that of task analysis. We conclude with a brief demonstration of this descriptive power by using it as an organizing framework of the evaluation of a collaborative virtual environment.

Keywords: Activity Theory, CHAT, Collaborative Virtual Environments.

Introduction

In contrast to the Western tradition of task analysis stands a parallel historical development, namely that of *activity theory*. For the purposes of this essay we are interested in introducing only one strand of this work, which is more fully described as *Cultural-Historical Activity Theory* (CHAT, though the terms AT and CHAT are generally used interchangeably). Our reasons for introducing the reader to this are several-fold. Firstly, there is something intrinsically interesting about considering an alternate aetiology and subsequent lineage of an independent line of research and reasoning concerning task analytic approaches to understanding the dynamics of work. Secondly, the collective nature of CHAT, derived from its Marxist roots, potentially offers a means of answering the now well established criticisms of task analysis and of human-computer interaction (HCI) as a whole (e.g. Bannon, 1991). Thirdly, CHAT finds a place for uncomfortable issues (at least to Western thinking) such as the roles of consciousness and motivation in human purposive activity. Clearly we have set ourselves a challenge in making a case for CHAT, but we also address the greater challenge for demonstrating the utility of CHAT in action. Here we will use an illustration drawn from a collaborative virtual reality (CVE) development project – DISCOVER and show how CHAT can define and organise the evaluation of the resulting CVE.

Origins

The aetiology of CHAT is complex, drawing upon as it does a number of different continental philosophical traditions, the most important of these were Marxism and Hegelian thought. Add to this its birth, during the early days of the Russian revolution, and we have a system of thought that is necessarily collectivist and socialist. A system, or body, of thought is a better description of *CHAT*, for it is not a theory, in the sense that it is not falsifiable or predictive in character - which are two of the hallmarks of scientific theory. Instead *CHAT* provides a strongly descriptive conceptual framework and vocabulary, directly attributable to the work of Lev Vygotski. Vygotski was a typical all-round genius living in the early twentieth century, who contributed to pedagogical thought, learning theory, the psychology of language and thought, developmental psychology and then promptly died of tuberculosis in his late thirties. His work outwith CHAT, particularly, in the domains of pedagogy, child psychology, and the psychology of language and thought remain standard texts and are still taught to undergraduates. A reasonably thorough treatment of his work could run to many volumes, so instead we will consider two or three of his key observations, before turning to the contribution of one of his students, Leont'ev.

Comment [TM1]:

Vygotski

There is no definitive biography of Vygotski but Alex Kozulin's introduction to the former's *Thought and Language* is generally regarded as to be the closest there is to one and upon which these brief notes draw heavily. A major theme of Vygotski's theoretical framework is that social interaction plays a fundamental role in the development of cognition. Vygotski (*ibid*) states: "Every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level; first, between people (inter-psychological) and then inside the child (intra-psychological). This applies equally to voluntary attention, to logical memory, and to the formation of concepts. All the higher functions originate as actual relationships between individuals." (*ibid*). A further aspect of Vygotski's thinking is the idea that the potential for cognitive development is limited to a certain time span which he calls the "zone of proximal development" (ZPD). For Vygotski this is the zone of proximal development, which he defines as,

> *The distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers.*

Vygotski, *ibid*

The spatial metaphor in the above quotation is mirrored in the real world as a zone or field to which an individual belongs which may be populated by experts, tools and other cultural artefacts. Vygotski's theory was an attempt to explain consciousness as the end product of socialization. For example, in the learning of language, our first utterances with peers or adults are for the purpose of communication but once mastered they become internalized and allow "inner speech".

Leont'ev

Leont'ev was one of Vygotski's students who, amongst other things, went on to develop a number of key concepts for what was to become activity theory. Unlike traditional task analysis, Leont'ev argued for the study of activity to be based on an understanding of the individuals' *object*, which is usually interpreted as *objectified motive* – motive made visible or tangible. This argument has three important consequences: firstly it allows us to identify uniquely a unit of analysis – the activity – by distinguishing between motivations. An activity is uniquely identified by its motive which is collective – the motivation of a (small) group. Secondly, it also betrays the psychological nature of activity theory. Both Vygotski and Leont'ev were psychologists and pedagogues and not designers or work study specialists. This, however, also introduces a consequential problem in that motivation cannot be observed directly; it can only be inferred. Thirdly, we can also introduce the concept of alienation (drawn from Marxism) wherein an individual's motivations are at odds with the collective's motivation. Activities are realised by way of an aggregation of *mediated actions*, which, in turn, are achieved by a series of low-level operations. This structure, however, is flexible and may change as a consequence of learning, context or both.

Comment [TM2]:

Comment [TM3]: is individual(s) singular or plural?

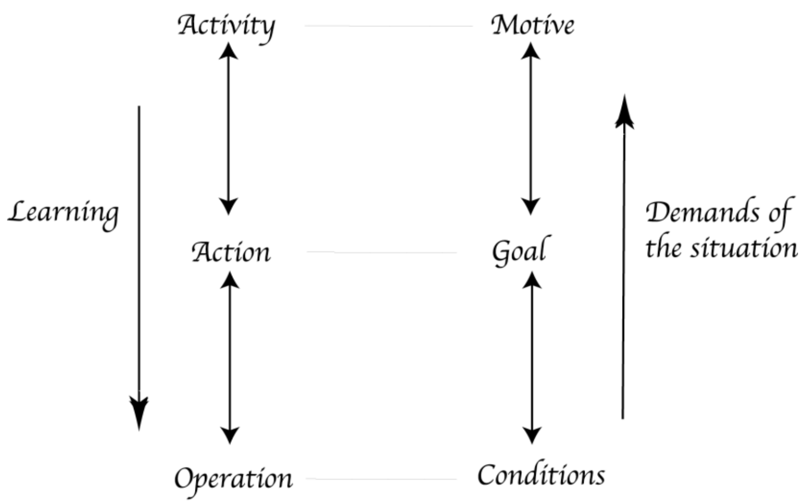

Figure 1 – The activity hierarchy

By way of example, consider the process of learning to use a complex interactive device such as a motorcar. The object of the activity is probably quite complex, ranging from and probably including the need to be able to drive because of work commitments; the need to attract the opposite sex; because of peer pressure; because an indulgent parent has given you a car or the need to participate in a robbery. The activity is realised by means of an aggregation of actions (i.e. obtain driving license; insure car; take driving lessons; learn the Highway Code; get a job to pay for the petrol and so on). These individual actions in their turn are realised by a set of operations – (i.e. get driving license application form, complete form, write out cheque for the license, send off license …). This, of course, is an incomplete, **static** description of the activity whereas humans are constantly learning with practice, so when first presented with the intricacies of the gear-stick (manual gear shift) it is likely that the process of disengaging the engine, shifting gear and re-engaging the engine are under conscious control (thus the action of changing gear is realised by the following operations - depress clutch, shift to the top left, release clutch). Thus the focus of attention is at the operations level but with practice attention will tend to slide down the hierarchy as the action becomes automatic. Over time actions become automatic and the activity itself is effectively demoted to that of an action – unless circumstances change. Such changes might include driving on the right (the British drive on the left) or change the make of motorcar or driving a lorry or being faced with the possibility of a collision. In such circumstances consciousness becomes refocused at the level demanded by the context.

Thus, this alternate formulation of the nature and structure of an activity is of interest for a number of reasons: firstly, this theory of activity, which has, at its heart, a hierarchical task-like structure. Secondly, it introduces the ideas of consciousness and motivation at the heart of the activity. Leont'ev offers a mechanism by which the focus (and locus) of consciousness moves between these various levels of abstraction – up and down the hierarchy depending on the demands of the context.

The role of Engeström in modern CHAT

> *To be able to analyze such complex interactions and relationships, a theoretical account of the constituent elements of the system under investigation is needed … Activity theory has a strong candidate for such a unit of analysis in the concept of object-oriented, collective and culturally mediated human activity.*

p.9 Engeström and Miettinen, 1999

In the late 1980s Engeström, a Finnish academic, has extended these ideas to include a model of human activity and methods for analysing activity and bringing about change in organisations in a manner which is reminiscent of participatory design (e.g. Engeström, 1987; Cole and Engeström, 1993; Engeström, 1995; 1999a; 2000). Engeström's work has been adopted and elaborated by many Scandinavian (e.g. Bardram, 1998a and 1998b), American (e.g. Nardi, 1996), Australian (e.g. Hassan, 1998) and British researchers (e.g. Blackler, 1993, 1994; Turner *et al.*, 1999, 2000, 2001a and 2001b). Engeström's account of activity theory is probably the dominant formulation in use in the study of information systems, HCI and CSCW research. In such research there is perhaps a greater focus on the role of activity *per se* rather than history and culture.

As we have already discussed, central to activity theory is the concept that all purposive human activity can be characterised by a triadic interaction between a *subject* (one or more people) and the group's *object* (or purpose) mediated by *artefacts* or tools. In activity theory terms, the subject is the individual or individuals carrying out the activity, the artefact is any tool or representation used in that activity, whether external or internal to the subject, and the object encompasses both the purpose of the activity and its product or output. Subsequent developments of activity theory by Engeström and others have added more elements to the original formulation. These are: *community* (all other groups with a stake in the activity), the *division of labour* (the horizontal and vertical divisions of responsibilities and power within the activity) and *praxis* (the formal and informal rules and norms governing the relations between the subjects and the wider community for the activity). These relationships are often represented by an activity triangle. Thus activities are social and collective in nature (please see Figure 2).

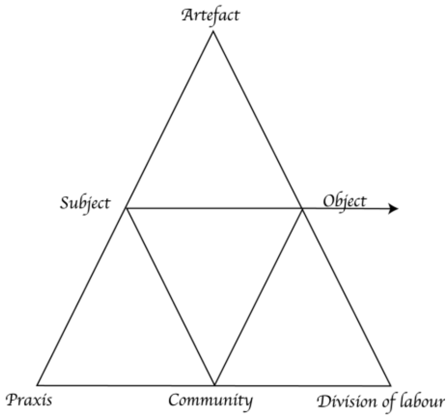

Figure 2: An activity triangle (schema)

The use of activity triangles is widespread in the activity theory literature but it must be remembered that this is only a partial representation of an activity. The triangle should be regarded as a nexus, existing as it does in a continuum of development and learning and in turn masking its internal structure of the individual *actions* by which it is carried out.

Activity theory is perhaps unique among accounts of work in placing such a strong emphasis on the role of collective learning. Vygotski's work on developmental learning has been a major influence on the thinking of Engeström, who has extended the idea to encompass collective learning which he has termed *expansive learning* (Engeström, 1987). Engeström has demonstrated the usefulness of expansive learning with its cycles of internalisation, questioning, reflection and externalisation in the development of activities in a variety of domains (Engeström, 1990; 1997; 1999). The drivers for these expansive cycles of learning and development are *contradictions* within and between activities. While this is something of a departure from Vygotski, it has proved particularly valuable to HCI and CSCW researchers (e.g. Holt and Morris, 1993; Nardi, 1996; Turner *et al.*, 2001a). We now consider contradictions in more detail.

Engeström's description of contradictions

Activities are dynamic entities, having their roots in earlier activities and bearing the seeds of their own successors. They are subject to transformation in the light of *contradictions*. Figure 3 is an illustration of an activity system (i.e. a group of related activities) complete with potential contradictions. Those contradictions found within a single node of an activity are described as *primary* contradictions. Primary contradictions might manifest as a faulty mediating artefact (e.g. bug-ridden software) or as heterogeneously composed subject group with, say, ill-matched training and skills.

The next category of contradictions is those that occur between nodes and these are described as *secondary* contradictions. In practice, this kind of contradiction can be understood in terms of breakdowns between actions or sets of actions that realise the activity. These actions are typically poly-motivated, i.e. the same action executed by different people for different reasons, or by the same person as a part of two separate activities and it is this poly-motivation which may be at the root of subsequent contradictions.

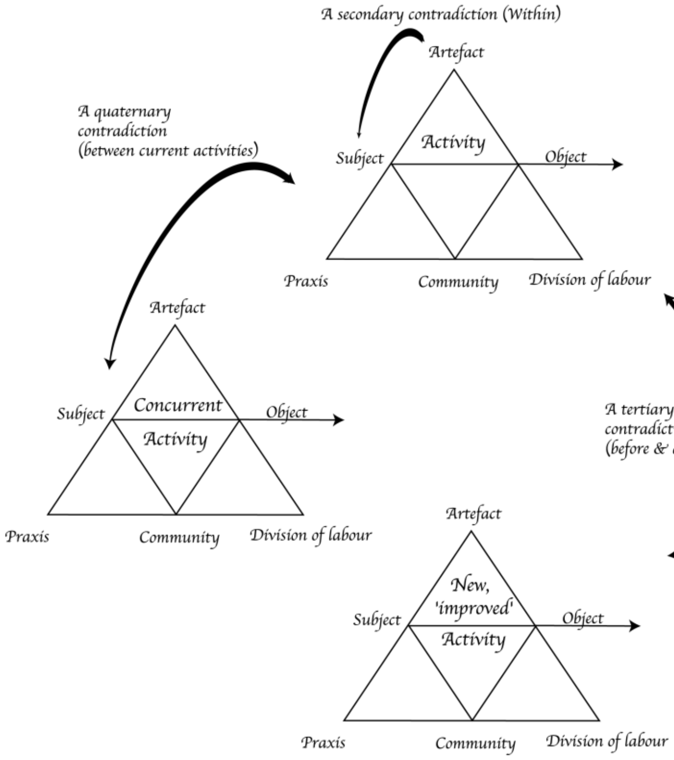

Figure 3 – An activity system and potential contradictions

Tertiary contradictions may be found when an activity is remodelled to take account of new motives or ways of working. Thus they occur between an existing activity and what is described as a 'culturally more advanced form' of that activity. A culturally more advanced activity is one which has arisen from the resolution of contradictions within an existing activity and may involve the creation of new working practices (praxis) or artefacts or division of responsibilities. Finally, those occurring between different co-existing or concurrent activities are described as *quaternary* contradictions. From this, it can be seen that a complex and continuing evolving web of contradictions may emerge. Primary and secondary contradictions in an activity may give rise to a new activity which in turn spawns a set of tertiary contradictions between it and the original activity and this is may be compounded by quaternary contradictions with co-existing activities.

Identifying contradictions

Having described the role of contradictions and how they may be classified, we now turn to how they are identified. Engeström (1999) describes contradictions as being identified by disturbances in the free running of an activity. Thus at the local level (e.g. local to an office or specific organisational division) these might include bottlenecks, varying folktales as to why a procedure is the way it is, differences of opinions as to the what, why and when of an activity. Engeström (2000) gives examples of such disturbances from a medical case study as including such things as mismatches between administrator forms; the uncertainty about the division of responsibilities between doctors and nurses and the sequencing of procedures. A further example is Holt's retrospective analysis of the contradictions in the activity systems operating in events leading up to the Challenger shuttle disaster. They concluded that fundamental contradictions in and between the activities underpinning the development, construction and deployment of the shuttle were ultimately responsible for the loss of the vehicle and crew (Holt and Morris, 1993).

Contradictions are distinguished from disturbances in that many disturbances may map onto a single contradiction. Thus disturbances are the visible manifestations of underlying contradictions. In practice this means that understanding the dynamics of the current work, making visible its nuances and identifying any disturbances therein are the necessary precursor to identifying contradictions.

From Description to Evaluation

Moving from this overview of activity theory, we now turn to a demonstration of activity theory in action. As we have said, activity theory offers a structured and structural account of an activity complete with its context and indications of internal dynamics and contradictions. These strengths can now be used to order the evaluation of a novel application *in situ*, namely, a collaborative virtual environments (CVE). We begin by describing the work, in this instance – training, the CVE is intended to support and then turn to our brief as regards the evaluation.

The Case study

The importance of safety-critical training in the maritime and offshore domains is recognised by all stakeholders in these industries, but is almost prohibitively expensive. Current methods require trainees to be co-located at a specialist training site, often equipped with costly physical simulators. The DISCOVER project aimed to provide a CVE based series of team training simulations which would dramatically reduce the need for senior mariners and oil rig workers to have to attend courses at specialist centres. While the system would be made available at such institutions, it could also be used over the Internet from offshore or on board ships. The consortium comprised four marine and offshore training organisations based in the UK, Norway, Denmark and Germany; virtual reality technology specialists, training standards bodies; and a number of interested employers and Napier University.

Current training

Space prevents a full treatment of safety-critical training at all of the training organisations so we shall confine ourselves to one example, which we shall call the Centre. At the Centre, the training scenarios are played out in a room adapted from a conventional lecture room. The 'bridge' area is found behind a screen in one corner of the room, and contains the ship's blueprints laid out on a table, alarm and control panels, communication devices and various reference manuals and a crew list. The other piece of simulation equipment is in the main body of the room. This comprises a set of four shelves rather resembling a large domestic tea-trolley each bearing the relevant blueprint plan for a four-deck section through the ship. Both these plans and those on the 'bridge' can be annotated with schematic depictions of hazard such as smoke, and are populated by miniature models of crew members who can be moved around, knocked over to simulate injury or death and so on. The 'trolley' can be seen in the figure below. The simulation is completed by an 'engine room', located in one of the tutor's offices down the corridor from the lecture room, and simply equipped with a pair of walkie-talkies and more blueprints.

A typical scenario at the Centre concerns a badly maintained ship taken over by the current crew at short notice, and carrying a hazardous cargo, which subsequently catches fire. A fire team is sent to investigate, and the situation is exacerbated by crew members being overcome by smoke, power failures, engine stoppages and sundry other hazards. Trainees form teams of the bridge party, the party dealing with the incident at first hand (working around the ‘trolley’) and the engine room. Other trainees act as non-participant observers. Tutor-trainee interaction is intense, relating both to the plot of the scenario and the team’s handling of it – tutors point out aspects which the team might have overlooked, hint at possible actions and generally keep the action running smoothly. As problems escalate, our observation confirms reports of the perceived realism of the session - the teams become very evidently engaged in the action, as can be seen in the figure, which is a still image from a training session. The figure in white is one of the trainers, the others trainees. They all can be seen to be using communication devices (cell phones and walkie-talkies). At the centre of the scene is the ‘trolley’ mock-up of a section through the ship. The trainer is in the act of moving some of the figures representing the remainder of the crew into the casualty position, a development that will be reported back to the bridge by the incident team leader. Once the action has run its course, a full debriefing takes place, comprising discussion and feedback about the teams’ actions, the real life scenario, and alternative approaches. Tutors take pains to ensure this is trainee led, and discussions are amplified by the tutor’s recall of particular incidents together with the incidents noted by observers.

Staff at the Centre have varying models of how the DISCOVER CVE might support their work. From the organisational point of view, it is hoped that the system will enable training to be delivered in a more flexible and economical manner, allowing skills to be acquired, practised and even assessed without the need for mariners to attend in person. This model requires an environment which is self-contained, supports all the different types of interaction described above, runs over the internet, and has the added value of simulating some conditions more realistically than current methods. Tutors would need the facility to modify events in the environment, as in current practice. In this view of the world, trainees interact with the environment, each other and any other role players inside the CVE, with the possible addition of video-conferencing for discussions, debriefing and tutor–trainee interaction. Another view expressed has been that the CVE would be a more-or-less direct substitute for the ‘tea trolley’ embodying the section through the ship, with the advantages of increased realism. Here trainees would remain physically co-located at the centre, and most interpersonal interaction would be outside the environment. It will be appreciated that the first of these alternatives is much more demanding, both in technical terms and in its implications for organisational change. Interviews with training staff indicate that the concept of remote delivery is seen as an interesting development with added potential for an enhanced degree of realism, the acquisition of new skills for themselves and additional business.

Figure 4 : Around the ‘trolley’ at the Centre. The man in white is a trainer, the other two characters are trainees (both of whom are ships’ masters). The tea trolley is the wooden structure between the men. Each shelf of the trolley represents a ship’s deck. The deck’s are populated by Subbuteo football figures representing members of the crew.

Comment [TM4]: Ship’s?

The envisaged DISCOVER training solution

The CVE itself was designed to run on standard, high-end, networked desktop PCs, the only special purpose equipment being high specification soundcards and audio headsets. The environment represented the interior of a ship (maritime version) or an offshore platform (offshore version). The users were present as avatars in the environment. Trainee avatars had abilities designed to mimic real action and interaction as closely as possible, and had access to small number of interactive objects such as fire-extinguishersfire extinguishers, alarm panels and, indeed, bodies. Communication again imitated the real world, being mediated through voice when in the same room, or telephone, walkie-talkie or public announcement (PA) when avatars were not co-present. Tutors were not embodied as avatars, but had the ability to teleport to any part of the environment, to see the location of trainees on a map or through a bird’s eye view, to track particular trainees and to modify the environment in limited ways, for example by setting fires or locking doors. It should be stressed that the environment was intended to support the training of emergency management and team coordination skills, rather than lower level skills. Figures 5 and 6 are screenshots taken from the maritime version of the DISCOVER CVE.

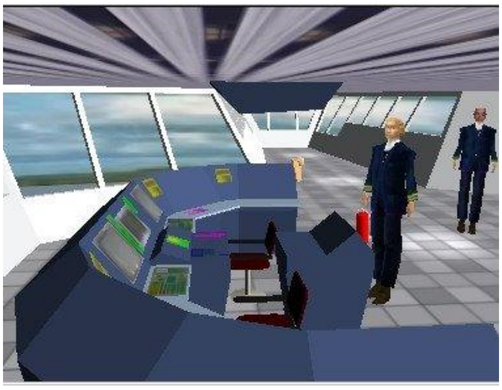

Figure 5: Avatars on the bridge

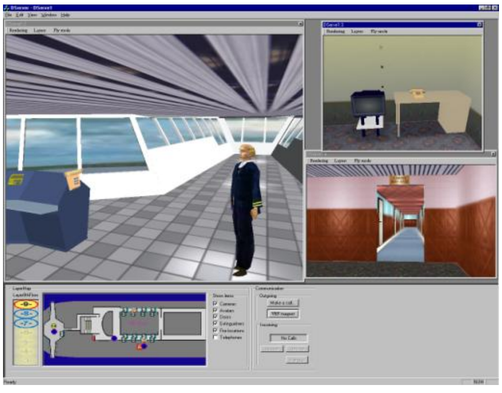

Figure 6: The tutor’s view of the bridge

The tasks of evaluation

As is usual in projects such as DISCOVER, we had defined the purpose of the evaluation in the project proposal in open and fairly vague terms, for example, does the CVE support the learning of key skills. In practice, what emerged from actually engaging in the project proper, was a need to address a wider series of issues. The first of these issues to emerge was one of acceptability – for both trainers and trainees. The second set of issues was the need to partition the evaluation itself logically, by adopting a suitable framework. We now address these in turn.

Comment [TM5]: series of wider issues?

Is the software acceptable?

We undertook a contradictions analysis of the current training activity to investigate the acceptability of the CVE to the trainers, trainees, training organisation and training standards bodies. The first step in this process was to identify disturbances in the current activity and between the current and the 'culturally more advanced' (new, improved) activity.

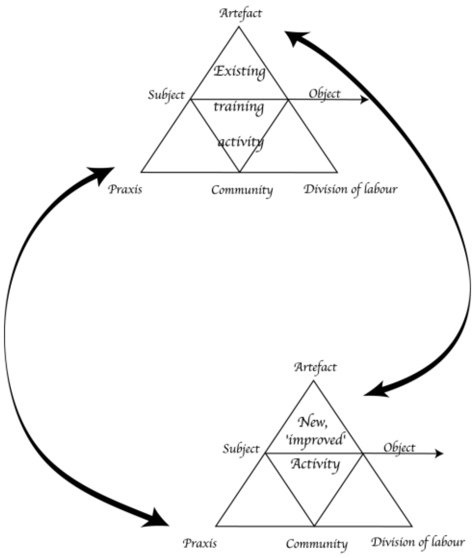

Figure 7: Potential contradictions between the existing and the new (improved) training activities

It should be stressed that the following discussion of a sample of the disturbances and potential contradictions we present are intended to be indicative only and are not intended to be either canonical or definitive.

Disturbances and potential contradictions within the current activity

1. Concerning the tea trolley. This prop had been proved to be very effective in supporting collaborative decision making but remained, well, just a tea trolley. There is a matter of image for the Centre. Resolving this contradiction - effective but *low-tech* - became one of the major thrusts of the DISCOVER project. Thus the challenge was to build a CVE with all of the flexibility, ease of use and effectiveness of the trolley. No easy task.
2. The problem with travel. Senior mariners are valuable people and had to travel to the Centre to be trained. Their absence was expensive for their employers, travel to and from the Centre was financially demanding as was their accommodation and so forth. Yet getting all of these people together in one place was the most effective of training them. This problem also generated a set of requirements of the DISCOVER system and was resolved by designing the CVE to operate in a distributed fashion over the Internet.

Potential contradictions between the current and new activities

3. The *Doom* problem. This problem was identified by the trainers themselves. They posed the question, "How can we seriously deliver safety-critically training using something which looks like a video-game?".

4. *Training the trainers* problem. The trainers are experienced senior mariners and expressly not computer people (their words) and they collectively expressed concerns about the need to train them in the use of the DISCOVER CVE. They also openly speculated about the consequences of introducing the technology for the structure of their jobs.

5. The identity problem. This problem was identified by one of the maritime standards bodies. They raised the issue of using DISCOVER in a distributed manner in places like South-East Asia and on the high seas – how could we guarantee the identity of the person engaged in training – given that they are collaborating remotely?

Comment [TM6]: Needs a trademark sign?

Each of these potential contradictions and many more were, in the main, were worked through in a series of what-if scenarios. Others were largely glossed over with the excuse of we will have to wait and see. We now turn to the structured evaluation of the CVE.

A structured evaluation of the CVE

We now turn to the structured evaluation of the CVE which we have based around the three level model of an activity as shown in figure 1.

Evaluating the purposive – activity - level

The purpose of the CVE was to support the teaching and learning of emergency management skills for offshore and maritime context. Related to this, it was essential that stakeholders should have confidence in the software as affording a means for such training, and trust that the skills learnt in the environment would be effective in real emergencies. Clearly, the evaluation of the fitness for purpose can only be undertaken with the participation of individuals from the community concerned. In one of the trials of early versions of the software, we had access to several maritime officers (including the Captain of a well-known passenger ship) who completed custom-designed questionnaire items about their confidence in the future use of the system as well as taking part in debriefing sessions.

More substantive evaluation for perceived fitness for purpose focussed firstly on data from the tutor sessions discussed below. Here data was collected through custom-designed questionnaire items, post-trial discussions and analysis of verbalisations and behaviour from the video record. As for pedagogic effectiveness, trials were planned with trainees in an employer organisation that incorporate realistic training scenarios with inbuilt checkpoints for the display of specific management behaviours at appropriate times. This will be complemented by observations based on the measures of team effectiveness derived by the TADMUS project (Salas and Cannon-Bowers and Salas, 1997) in their research into training for decision making under stress, and on the deeper aspects of pedagogy in the Laurillard model (Laurillard, 1993). Finally, DISCOVER must receive the seal of approval from industry validating bodies.

However, despite this detailed planning it proved to be impractical to run rigorous comparative trials of DISCOVER against conventional training (because of the restricted availability of trainees and the related difficulty of ensuring matched groups). Still less will it be possible to 'prove' the effectiveness of DISCOVER training in genuine emergencies. It remains the case, that at the current state of knowledge, the verification of the transfer of VR-based training into the real world is very much an active issue for research. (Caird, 1996 explores these issues in some detail).

Action level evaluation

Here the focus of evaluation is how effectively are actors embodied *in* the environment and how effectively they can collaborate *through* the environment. In addition, we are concerned with evaluating the related issues of perceptions of fidelity, presence and engagement.

Trainees in the DISCOVER environment needed to be able to find each other, to communicate by appropriate means with fellow trainees and tutors, to monitor what others were doing and to interact with various items in the environment, for example to pick up a body (an avatar) overcome by smoke. Tutors had to be able to gather sufficient information from monitoring activity in the CVE to provide guidance and post-training feedback, to communicate with trainees and to modify interactive objects in the CVE such as the location of fires. It had also been stressed by all stakeholders from training and employer organisations that the CVE must be extremely realistic and imbue a strong sense of presence if it was to be considered fit for its purpose of providing training. This was for two reasons. Firstly, existing physical ship simulators are extremely close to their sea-going equivalents, so much so that officers undergoing simulator training can be dismissed should they run the simulator aground. Secondly, one of the key elements in emergency management training is engagement in the emergency scenario, and consequently the experience of a suitable degree of stress.

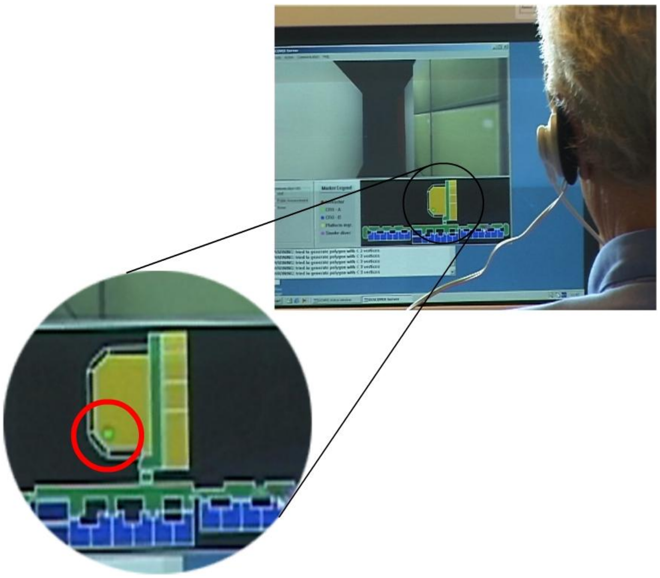

Figure 8 - A trainer deciding which way round he is facing. This is because his represented by a flashing circle.

Comment [TM7]: |What is?

Comment [TM8]: He is?

Here the choice of techniques was constrained by the limited range of ready-made tools for evaluating aspects of collaboration in virtual environments, and again by the availability of subjects. Aspects of communication and coordination (primarily, being able to see, hear and address other users) were evaluated in parallel with the ergonomic elements in the very early trials described above. Once the software was reasonably stable and more co-working features had been added, more complex trials were carried out.

Comment [TM9]: Extremely long paragraph – have attempted to split it at natural points

As before, largely proxy subjects were used to identify the most immediate issues concerning affordances for embodiment and communication. They undertook structured to include (i) the type of collaborative tasks undertaken in a realistic training situation and (ii) the underlying collaborative actions identified the COVEN hierarchical task analysis (COVEN, 1998). Short post-use questionnaires were administered using items derived from the task analysis.

Comment [TM10]: They undertook tests/tasks structured?

Comment [TM11]: By the COVEN hierarchical task analysis (approach?)

For users adopting the role of tutor, an additional set of tasks and questionnaire items was derived from Laurillard’s (1993) model of teaching and learning. At this level we did not seek to address the efficacy of any teaching or learning, but rather the support of the environment for such pedagogic actions as setting/modifying task goals, monitoring trainees and giving feedback. Again, observers monitored the progress, or, occasionally, lack of progress, of the scenario, supported by checklists mirroring the questionnaire content.

Finally, issues of fidelity and presence were also covered. Initially this was through a short series of items in the post-use questionnaire and observers' checklist, again adapted from VRUSE. The final version of the software was evaluated with experienced tutors from one of the training organisations involved in the project. (Evaluation techniques had been planned for trials with 'real' trainees, but in the event personnel could not be made available. This work continues outside the scope of the project at one of the training organisations). Tutors undertook a realistic training scenario, authored by one of the training organisations. They took turns to play tutor and trainee roles. This time the NASA ITQ questionnaire (a measure of immersive tendencies, Witmer and Singer, 1998) was administered before the trial started, followed up by a questionnaire instrument incorporating the collaborative and pedagogic aspects as before, coupled with the NASA PQ – the counterpart to the ITQ which aims to measure presence. These trials were videotaped for further analysis of evaluation data.

Evaluating the operational level

The aspects of the CVE to be evaluated at this level are those concerned with the ergonomics and usability of the means provided to interact with the CVE. These include the now standard range of GUI controls, as well as input and output devices. Aspects to be considered are their perceptibility, ease of operation, provision of feedback and, in general, the list of low level usability heuristics to be found in any textbook. Here we were concerned, *inter alia*, with affordances of such features as the push-buttons provided to activate virtual communication devices such as the phone and walkie-talkie and the use of the mouse click as a means of opening doors, setting off fire extinguishers and generally activating objects. The design of these had been a subject of much debate as to whether, for realism, a phone should have the usual set of buttons reproduced virtually, or if users would find a dialogue box more convenient. We also needed to evaluate the basic input devices, such as the mouse, for moving through the environment (employers were keen that the system should run on a standard workshop and peripherals), and of the headsets used for verbal communication.

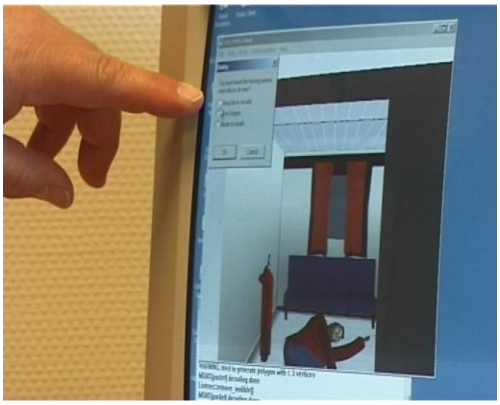

Figure 9 - Examining a dialogue box

The overall emphasis in the choice and construction of techniques for level 1 was to obtain basic usability data with minimal consumption of analyst and user resources. These affordances were primarily investigated through user trials, starting from the earliest versions of the software. Early trials employed very largely 'proxy' subjects who represented the eventual user population as closely as possible in terms of relevant background skills and experience. This allowed us to conserve the scarce resource of 'real' users for both more polished versions of the software and fitness for purpose issues. Subjects undertook realistic single-user and collaborative tasks matched to the functionality of the software version under review, monitored by observers. Figure 9 shows a user contemplating a dialogue box. With later trials the main evaluation focus shifted to the action and operation levels, but usability continued to receive some attention. Post-trial questionnaires were compiled and administered, adapting usability items from standard usability instruments and VRUSE (Kalawsky, 1999), and guided by the insights in Kaur, Maiden and Sutcliffe (1997). Although the custom built questionnaire did not now have the strong validation of its parents, the questions could be tailored to the particular context of the DISCOVER CVE while keeping the overall instrument to a manageable length. Observers augmented the self-report data. The trials were supplemented by usability inspections structured by standard heuristics.

In the event, most usability problems were identified by a initial, quick expert check of the interface, but the other techniques adopted were able to provide substantive data to back up these observations.

Discussion and conclusion

This chapter began with a brief description of activity theory and in doing so we have stressed its power of description. We then attempted to demonstrate the utility of an activity theoretic framework in organising the evaluation of a CVE. It remains the judgement of the DISCOVER project team that the framework proved to be valuable but we recognise that it needs further development and validation.

As for CHAT more generally, we are not alone is seeing its potential as a descriptive framework for disciplines such as human-computer interaction and computer supported cooperative working but this potential remains largely unrealised. It undoubtedly offer a different and hopefully complementary take on human purposive behaviour to that of mainstream task analysis. However in reviewing the application of CHAT in related domains we have found very little evidence for its prospective use – as opposed to retrospective descriptions of the work, situation and device on study. Thus, apart from CHAT's relatively obscurity, the key challenge for activity theorists it to map its rich descriptive power onto a prospective methodological framework.

Acknowledgements

Special thanks to Dr. Susan Turner for leading the evaluation workpackage and for letting us rifle her evaluation deliverable for *bon mots*. Finally thanks to Oli Mival for capturing the evaluation footage (or is it meterage?) and for the screenshots. We gratefully acknowledge the contributions of our colleagues on the DISCOVER project in providing the sites and subjects for the fieldwork herein described and in developing the DISCOVER software. The project was financially supported by the EU ESPRIT programme.

References

Bannon, L.J. (1991) From Human Factors to Human Actors. In J. Greenbaum & M. Kyng (Eds.) *Design at work: Cooperative Design of Computer Systems*. Hillsdale: Lawrence Erlbaum Associates, 25-44.

Bardram, J. E. (1997) Plans as situated action: an activity theory approach to workflow systems. In J. A. Hughes, W. Prinz, T. Rodden and K.Schmidt (Eds.) *The Proceedings of the Fifth European Conference on Cosmputer Supported Cooperative Work,* Kluwer Academic Publishers, 17-32.

Bardram, J. E. (1998a) Scenario-based Design of Cooperative Systems. In F. Darses and P. Zaraté (Eds.) *The Proceedings of COOP '98*, INRIA, 20-31.

Bardram, J. E. (1998b) Designing for the dynamics of cooperative activities. In *The Proceedings of the CSCW'98*, ACM, 89-98.

Blackler, F. (1993) Knowledge and the theory of Organizations: Organizations as Activity Systems and the reframing of management, *Journal of Management Studies*, **30(6)**, 863-884.

Blackler, F. (1994) Post(-)modern organizations: understanding how CSCW affects organizations, *Journal of Information Technology*, **9**, 129-136

Caird, J. K. (1996). Persistent issues in the application of virtual environment systems to training. *IEEE: Human Interaction with Complex Systems,* **3**,124-132.

Cannon-Bowers, J. A. and Salas, E. (1998) *Decision making under stress*, APA, Washington.

Cole and Engeström (1993) A cultural-historical approach to distributed cognition. In G. Salomon (ed.) *Distributed Cognitions – Psychological and educational considerations*, Cambridge University Press.

Cole, M. and Engeström, Y. (1993) A cultural-historical approach to distributed cognition. In G. Salomon (ed.) Distributed Cognitions – Psychological and educational considerations, Cambridge University Press, 00-01.

Engeström, Y. (1987) Learning by expanding: An activity-theoretical approach to developmental research. Helsinki: Orienta-Konsultit.

Engeström, Y. (1995) Objects, contradictions and collaboration in medical cognition: an activity-theoretical perspective, Artificial Intelligence in Medicine, 7, 395-412.

Engeström, Y. (1996) Developmental studies of work as a testbench of activity theory: the case of primary care medical practice. In S. Chaiklen and J. Lave (Eds.) Understanding Practice: perspectives on activity and context, Cambridge University Press, 64-103. First published in 1993.

Engeström, Y. (1998) The tensions of judging: Handling cases of driving under the influence of alcohol in Finland and California, In Y. Engeström and D. Middleton (Eds.) Cognition and Communication at work, Cambridge University Press, 00-01.

Engeström, Y. (1999) Expansive visibilization of work: an activity theoretic perspective, CSCW, 8(1-2), 63-93.

Hassan, H. (1998) Activity theory: a Basis for the Contextual Study of Information Systems on Organisations. In H. Hasan, E. Gould and P. Hyland (Eds) Information Systems and Activity Theory: Tools in Context, University of Wollongong Press, 19-38.

Holt, G. R. and Morris, A. W. (1993) Activity Theory and the Analysis of Organizations, *Human Organization*, **52(1)**, 97-109.

Johnson, C. (1999) Evaluating the Contribution of Desktop VR for Safety-Critical Applications, *Proceedings of SAFECOMP'99,* Springer Verlag Lecture Notes in Computer Science

Kalawsky, R. S. (1999). VRUSE - a computerised diagnostic tool: for usability evaluation of virtual/synthetic environment systems. *Applied Ergonomics*, **30**, 11-25.

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

Kaur, K., Maiden, N. and Sutcliffe, A. (1997) Interacting with virtual environments, an evaluation of a model of interaction. *Interacting with Computers, Special Issue on VR,* **11**, 403-426.

Kuutti, K. Activity theory and its applications in information systems research and design. In Information Systems Research Arena of the 90's. H.-E. Nissen, H.K. Klein, and R. Hirschheim, Eds., Elsevier, Amsterdam, 1991.

Kuutti, K. and Arvonen, T. (1992) Identifying Potential CSCW Applications by Means of Activity Theory Concepts: A Case Example. In J. Turner and R. Krauts (Eds.) *The Proceedings of CSCW '92*, ACM Press, 233-240.

Laurillard, D. (1993) *Rethinking University Teaching*, Routledge, London

Nardi, B. (1996) Some reflections on the application of activity theory. In B. Nardi (Ed.) Context and Consciousness, MIT Press, Cambridge, MA, 00-01.

Suchman, L.A. (1987) *Plans and Situated Actions*. Cambridge: CUP.

Turner, P. and Turner, S. (2001a) A web of contradictions, *Interacting with Computers*, **14(1**),1-14.

Turner, P. and Turner, S. (2001b) Describing Team Work with Activity Theory, *Cognition, Technology and Work,* **3(3)**, 127-139.

Turner, P., Turner, S. and Horton, J. (1999) From description to requirements: an activity theoretic perspective, *Proceedings of Group'99*, ACM Press.

Vygotski, L. (1986) *Thought and Language*. MIT Press, Cambridge, MA. Translated and edited by Alex Kozulin.

Witmer, B.G. and Singer, M.J. (1998) Measuring Presence in Virtual Environments, a Presence Questionnaire. *Presence*, **7(3),** 225-240