PETRA: Participatory Evaluation Through Redesign and Analysis*

CSRP No. 75, March 1995

Susi Ross **Magnus Ramage**† **Yvonne Rogers**
School of Cognitive and Computing Sciences
University of Sussex, Brighton BN1 9QH, U.K.
Phone +44 127 678524, Fax +44 127 671 20.
susi,yvonner@cogs.susx.ac.uk, magnus@comp.lancs.ac.uk

Abstract

Compared with single user-computer interactions, evaluating CSCW is difficult.

specific quantitative measures? Alternatively, is it possible or desirable to take the middle-road; achieving a balance between detail and generality?

The approach adopted will obviously depend on the goals of the evaluation. If the objective is to analyse the working practices and the nature of collaboration in a given organisational context, to inform system design, then the use of social methods such as ethnography would seem appropriate. Alternatively, if the aim is to analyse the usability of, say, World Wide Web browser tools for different user groups, then perhaps a cognitive task analysis would be more appropriate. In many situations however, it is not easy to determine the objectives of the evaluation because the purpose of the collaborative tool is not clear. For example, videoconferencing allows participants to see images of each other when talking at a distance. Given the ease with which people communicate via telephone, one might reasonably ask if there is any added value in being able to see distant callers in this format: much anecdotal evidence and the few empirical studies carried out suggest it is minimal and can even have adverse affects (Harper & Carter, 1994; Pagani & Mackay, 1993). This can be viewed as an example of the 'solution looking for a problem' phenomenon: potential benefits of the prototype tool are unknown, and useful contexts have yet to be discovered. Compared with evaluating single-user systems that have specific objectives (e.g. evaluating the new functionality of an update version), and for which benchmark tests can be relatively easily devised, evaluating CSCW systems is proving to be much more problematic.

Given our current lack of understanding of how best to utilise and benefit from the new and largely unfamiliar generation of CSCW systems that is emerging (e.g. video-links, shared drawing, writing and editing tools, shared repositories) - at work, home and elsewhere - we suggest there needs to be much more emphasis on determining their usefulness in different contexts. In particular, we need to develop evaluation methodologies that can show how the various tools are able (or not) to support collaboration, as well as provide HCI-style usability indicators (e.g. easy to learn, easy to use) appropriate for assessing groups of users using the CSCW tool together.

The aim of this paper is to begin addressing this situation. We start by advocating eclecticism, whereby different methods and theoretical frameworks are combined. Such an approach should allow for a more reflexive analysis, between different theoretical concerns and practical design issues, and in doing so force us to explore and make explicit many of our assumptions about the CSCW tools and the nature of the collaborative activities that they are intended to support. To this end we have developed a multi-perspective framework, called PETRA, which stands for 'Participatory Evaluation Through Redesign and Analysis'. Essentially, the framework brings together a theoretically-driven 'evaluator's perspective', based on a combination of theories that are concerned with collaboration (evaluation through analysis, or ETA), and a practically-based, user-focused 'participant's perspective', drawing from heuristic evaluation techniques and participatory design (evaluation through redesign, or ETR).[1] The primary objective in developing this framework was to obtain a detailed understanding of a collaborative activity whilst also considering how to inform the re-design of the interface for a CSCW tool. In particular, the former was intended to provide theoretical insight into the social and cognitive issues that are important for shared understanding when working together at different sites, whilst the latter was intended to provide an understanding of the system and interface attributes desirable or neccessary to successfully support the former.

In the next section we discuss the multiplicity of theory, method and perspective when evaluating CSCW; and we introduce the different parts of our PETRA framework. In Section 3 we illustrate PETRA in practice, describing how we used multiplicity in our study. Sections 4 and

[1] The participatory 'P' signifies the inclusion of both participants' and evaluators' opinions, and of multiple perspectives within the framework.

5 present findings; discussion and conclusions follow in Section 6.

2 Multiple methods, perspectives and theories

The diverse fields informing CSCW (eg. HCI, sociology, software engineering, psychology, cognitive science, management studies) bring with them a rich variety of evaluation methodologies, perspectives and

perspective methods are more suited to evaluating aspects of the tool. Yet evaluating activity or tool alone results in an incomplete picture. By recognising the interdependence of collaborative activity and supporting tool we can apply both evaluator and participant methodologies within the same study. This hopefully yields *complementary* and extensive coverage of the collaboration, yet without requiring too much time or resources.

In sum, the PETRA framework integrates multiple theories, methods, and perspectives. In doing so it incorporates the expertise of both participant and evaluator; combines theory and practical experience; and integrates use, iterative design and evaluation. How we used the PETRA framework in practice is described next.

3 PE RA illustrated: our analytic framework

Here we describe one particular instantiation of PETRA, devised to suit not only the activity (collaborative writing) and the tool (ShrEdit) in question, but also our goals and resources. The 'evaluation-specific' goals were to evaluate the efficacy of ShrEdit as a tool for the support of distributed computer-based collaborative writing; and to investigate some of the processes involved in the activity of collaborative writing via various mediums. We were also interested in testing out the PETRA approach; in developing some of the theoretical concepts underlying the evaluators' perspective; and in exploring the potential for 'participatory redesign' as a part of any evaluative method. Our resources were limited: three users; two evaluators; no budget; three machines in distributed locations; one video camera; and one tape recorder. We briefly describe the activity and tool next: the remainder of the section discusses the practical activities and analytic framework we devised around these constraints.

Within CSCW, collaborative writing is concerned with the study of multi-author writing processes, and the creation of group editors which support this process. Co-authoring is an everyday occurrence in numerous paper-based tasks, and is, potentially, the most common kind of computer-supported collaborative work. Computer-based collaborative work offers the advantage of connecting multiple distributed participants (often synchronously), thus opening new opportunities for people to work together where before it was too time-consuming, expensive or cumbersome. Here there are no precedents, and even more need for effective evaluation of the resources devised to exploit such opportunities. What do we know about collaborative writing in a face-to-face context, and will the same mechanisms be relevant to a computer-based setting? From paper-based collaborative writing we need to know *who* is involved in the writing process; *what* kind of document they are creating; *why* they are co-authoring; and *how* the process is structured. Then we must consider if it is either neccessary or desirable to recreate such mechanisms in a computer-based setting (Beck, 1994), and if the essentials of interaction are not altered by the very medium being used. We hope to address some of these issues in our study, particularly through our use of the evaluator-perspective. It should be emphasised that we use collaborative writing as an exemplar in our study of evaluation in CSCW.

system is used to prevent conflict, and ensure coordination: text being changed by one user is 'locked' (indicated by a padlock symbol) so that no-one else has simultaneous editing access to it. Both public and private windows are available, allowing work to be developed privately before being inserted into the shared document. Awareness of other users' actions is provided by the status panel, which allows a user to find the current editing location of another user, or to 'track' (follow) another's actions through the document. However, it is not otherwise possible to find out what another user is doing: in particular, to see who is entering a particular piece of text. We wished to evaluate the usefulness of ShrEdit as a collaborative text editor: in particular, we were interested in the attempted provision of awareness information. Were ShrEdit's mechanisms successful or sufficient to support group awareness? If not, why not, and what might be required instead?

Evaluating the activity: shared understanding through different media

We decided to focus our evaluator's analysis (ETA) of the collaborative activity on shared understanding through different communicative media (talk, paper and computer-based). By shared understanding we mean the way in which two or more people relate their common background and experiences to understand collectively what each is talking about. This concept has been characterised extensively in the literature as being central to communication, by social psychologists, cognitive scientists, sociolinguists and sociologists alike, under various guises including mutual knowledge, common ground, social organisation and intersubjectivity (Clark & Brennan, 1991; Garfinkel, 1967; Hutchins & Klausen, 1992; Krauss & Fussell, 1991). A major concern has been to explicate the mechanisms by which speakers establish and maintain shared understanding during conversations. It has been proposed that speakers achieve this by formulating their contributions in relation to their awareness of what the other persons know and do not know. Moreover, such common ground is never static, but has to be continuously updated moment by moment and from context to context.

In outlining the different mechanisms involved in shared understanding (e.g. conversational acts, rhetorical devices, repair strategies and non-verbal behaviour like gaze, nodding) some researchers have also considered how they change when different media are used. For example, Clark and Brennan (1991) summarise the various constraints that a medium can impose on communication between people. These include whether the media allow co-presence, co-visibility, co-temporality, simultaneity and revisability of messages sent. They argue that the presence or absence of these factors affect the mutual achievement of common grounding in different ways. Whereas in face-to-face situations all of the above are possible, only a restricted number of these are available with video-conferencing, telephone, email and letters. Consequently, Clark and Brennan argue that when a medium lacks one or more of these characteristics it can require the participants to use alternative grounding mechanisms, which have different costs associated with them. For example, when contextual cues are missing in communication (e.g. in email conversations) the costs can be higher; misunderstandings can arise requiring the use of repair mechanisms to maintain shared understanding.

In our study we were interested in the nature of any additional costs incurred when shared understanding is mediated through a computer-supported collaborative tool and an adjoining telephone compared with face-to-face settings utilising shared paper-based resources (i.e. notes, books). To characterise and analyse the different costs, we used notions from distributed cognition (see Rogers and Ellis (1994) for an introduction) and ideas from Heidegger.

Winograd and Flores (1986, p.77)) that occured in the different settings.

Previously, the distributed cognition approach has been used to analyse distributed working in a number of different environments - mainly 'control room' situations such as ship navigation (Hutchins, 1990), aircraft piloting (Hutchins & Klausen, 1992), air traffic control (Halverson, 1994) and offices such as an engineering design firm (Rogers, 1992) and hospital administration (Rogers & Ellis, 1994). Similarly, the use of breakdown analysis has been used to analyse various CSCW activities (Sharples et al., 1993; Urquijo et al., 1993). Hence, our analytic tools have already been used in various CSCW contexts. Our intention here was to use this particular combination to understand better the different mechanisms used in shared understanding when supported by different media. In particular, we wanted to find out how shared understanding develops when writing collaboratively. When a document is being constructed, we might argue that there is not just the document itself being created, but also a set of individual understandings of the document (or 'representations' in distributed cognition terms), which combine together to form a collective understanding. Specifically, we wanted to address the question of how the transition from a paper-based to a computer-based workspace, and from face-to-face dialogue to that over the telephone and through the shared work space, affected the development of shared understanding in the collaborative writing process.

Evaluating the tool: PD and the playschool

We argued above for the integration of users and evaluation, and evaluation and re-design. Our PD-derived ETR methodology allows us to incorporate both the direct experiences of the users, and the redesign process itself within the evaluation framework, giving us a strong participants' perspective. What is important is not so much the details of the methodology used but that it allow both a direct role for participants

pens; crayons; pencils; highlighters; ruler; scissors; glue; eraser; sellotape; stapler; blue tack; overhead transparencies and coloured pens; assorted coloured sticky labels; and Post-It notes. (We thought the pre-formed materials used in Pictive too restricting, and used only 'free-form' resources). These materials were available on a shared workspace table, with the participants seated around it. One evaluator acted as facilitator during the session, mainly to encourage the participants to use the materials, and to answer specific queries about system capabilities or CSCW research. The other operated the video camera, and logged interesting events. The final set-up and atmosphere was designed to enable effective exploration of usability problems through accessible and non-intimidating design materials.

There are interesting methodological issues here which merit further investigation. Should evaluators participate, or facilitate, and to what extent? How can the video diary be exploited to enhance the evaluation and redesign process? To what extent is such a Playschool dependent on the personalities of the group involved? For successful and integrated evaluation of design we believe that users should be allowed to express for themselves, both verbally and visually, *what they* found problematic; *why they* think that is; and *how they* think such problems could be solved. Undoubtedly there are many methods and techniques which can support this process, of which our Playschool session is only one example.

The study

A three-part study was conducted, using the same three participants in each part: they are referred to here as Charles, Simon and Nigel (not their real names). All three participants were students on the same Masters course, taking a class in Artificial Neural Networks.

In the first part, the participants collaboratively created a document in a face-to-face, paper-

(books and lecture notes) the only knowledge other participants had of their consultation was by explicit spoken or written references.

Much of the shared understanding that occured took place via social coordination over the shared telephone link. Compared with the face-to-face session, where talking was continuous, there were long periods of silence (up to 90 seconds) while all the participants were focussing on using ShrEdit. An example of shared understanding occuring through telephone mediation is as follows: Simon raises a question about how the current topic-of-composition works. His understanding is confirmed by Nigel, and then strengthened by Charles, leading to a stronger understanding for the group as a whole:

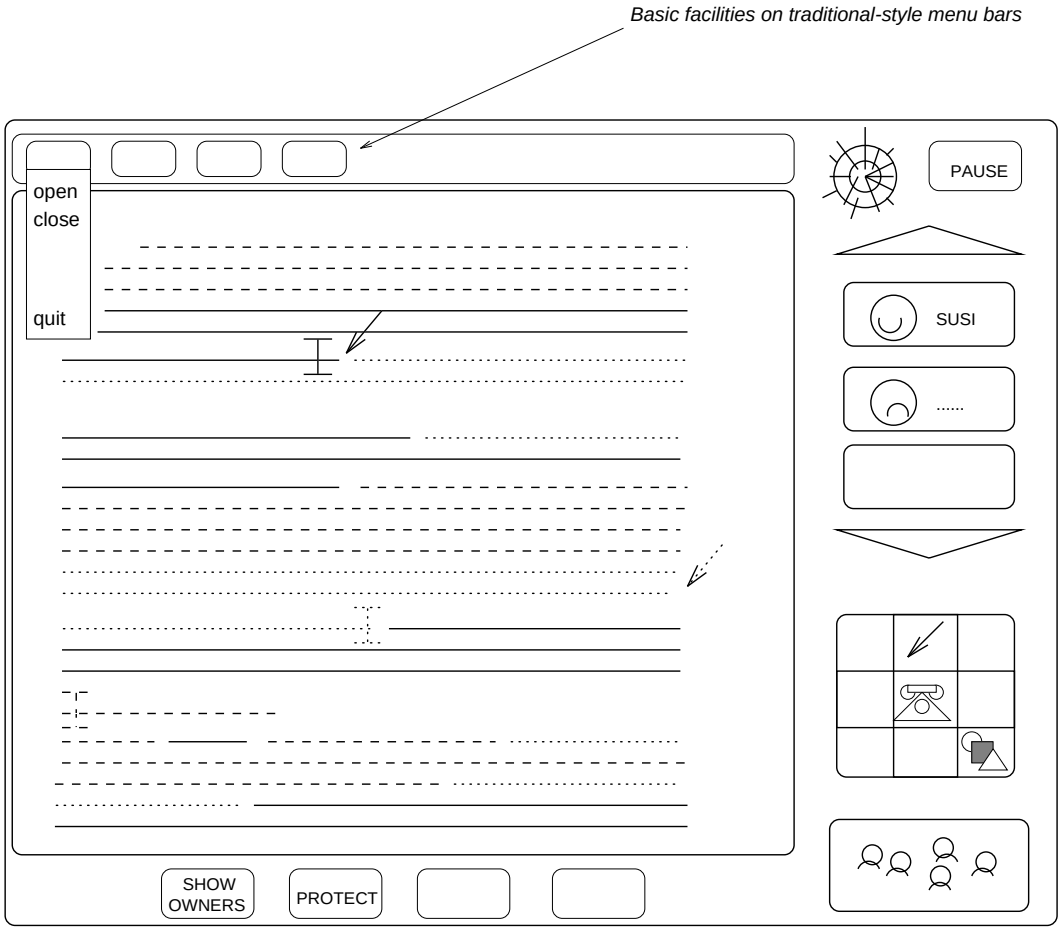
S: I'm just - just a general thing about the back propagation itself, right? Em, do you - you compute the errors on the outputs, right?
N: Yeah
S: And then you compute the errors on the input units. Oh no, it's all right, it's all right
C: What you do is you take - what you do is you calculate the error from the, from the output erm for the target and then
S: Yeah, I've got that
C: And then from that, you send that back
S: The error?
C: The error from the target
S: Right
C: Back into the understanding (generalati)99or101.(or)-1999.20what

writing did go on in the ShrEdit session, and of a much higher quality, suggesting that at least technologically-mediated collaborative writing leads to better writing, if not necessarily a higher degree of shared understanding.

5 Findings: redesigning the tool

This section describes the findings from our evaluation of ShrEdit as supporting tool in Part 3 of the study, using the Playschool ETR method. The Playschool session produces two main sources of evaluative information: the designs produced by the participants, which expressmediat8s20.41

COLOUR:
a colour is assigned to each participant when they sign on



> 'Is someone doing something?' (Charles, 02.50);
> 'Who's typing? Charles?... Someone's typing' (Simon, 05.15);
> 'is somebody on - Nigel where are you? Are you on...' (Simon, 09.00);
> 'Frustrating! there's no clue as to where you actually are' (Charles, 09.55).

Solving this problem was the main issue tackled during the redesign session. Each editing conflict was complicated by the lack of information as to who was involved. Without a speech-based communication channel to sort out such problems, we believe it would have been impossible to use ShrEdit with any efficacy at all. Colour was the key resource proposed in dealing with the problem of providing awareness information for multiple users:

> C: COLOUR! Colour highlighting! Give everyone a colour when they sign onto a session, and then all they type will be in that colour. Cos then you can see who owns what piece of text, and you can see where people are. (Playschool log)

Highlighting and cursors, as well as text, would all vary in colour with the user. The need to see each person's cursor was considered crucial, as currently only your own cursor is visible, giving others' editing changes an uncanny disembodied feel! The participants proposed a combination of a 'normal' cursor (perhaps with face attached) and a 'big pointer', to attract attention (Stefik et al., 1987). Text currently being worked on would also be highlighted, to cut down on accidental edit conflicts. Locked text could also be designated as such in the appropriate colour.

To support awareness of who is currently using the system, the participants suggested a side-strip giving faces, names, email, and project details of each group member, which would also indicate their 'active', 'be-right-back', or 'missing' status. This strip would then be used to activate various tools: email, 'talk', or 'attract' tools would be dragged to it, and dropped onto the desired head. Finally, it is worth noting that the participants themselves discussed the need for user control of all these facilities - coloured text, flashes and sound could all be turned on or off at will.

Communication

This is not usually an issue in a single-user context. In contrast, people engaged in collaborative work (even face-to-face, paper-based work) need to communicate with each other during the process. Of interest, therefore, is what processes are involved, and how best to support them? We see two main kinds of communication: peripheral and explicit. Peripheral or background communication occurs as a side-effect of something else, yet also serves to enhance awareness of current events. Explicit communication is where one user enters discussion with any or all of the others. Both forms need supporting. The question is, should it be verbal or text-based; private or public; integrated or an add-on application?

In our study regular use was made of the telephone conference-call link, an add-on to ShrEdit's basic functionality. Much conversation time was spent discussing application problems, yet the rest of on-phone conversation was spent in discussion of the structure and content of the document, and in organising the group. The participants explored various communication-channel possibilities during the ETR session, but a built-in microphone was felt to be the best option. It should have an on-off capability, a way of attracting attention, and perhaps be mounted on a head-set. A text-based message window was thought to be too slow, and cumbersome if several other documents were already open. However an additional, private talk window was considered useful, to allow private collaboration between selected group members on an ad-hoc basis. The ability to turn both microphone and speakers on and off was felt important to ensure both privacy and user control.

Focus

In single user-computer interactions the computer simply flashes or beeps to attract the user's attention. But in a world of multiple users - all doing different things at the same time to the same document - how do you ensure that each user is aware of changes without overwhelming their eyes and ears with colour, motion and noise? How do you attract and direct the user's attention, or enable each group member to attract the notice of selected others? Moran and Anderson (1990) discuss the idea of *mundane* technology: technology that is ever-available in the background, but never impinges unduly on the user's consciousness. This becomes a much harder task when multiple active users are involved. This category is closely related to both feedback and awareness, and to communciation: in giving feedback the computer will need to draw the users attention to some outcome or other; and users will need to attract others attention before communication can take place.

The participants did have problems attracting each other's attention: ShrEdit has no in-built way of doing so, nor did they develop any social protocols to deal with the situation (such as typing a message on-screen). Simply providing an integral communication channel does not solve this problem, as a suitable means of attracting attention must also be found. The participants suggested that they should be able to direct the attention of other group members to a particular item of text, for comment or editing. Whilst one advocated flashing text, others thought this too invasive and annoying. A flashing alarm symbol with accompanying auditory feedback and message identifying the caller was finally agreed upon[2].

Coordination, ownership and control

In a single-user application, the user retains control, and is sole owner of their work. In CSCW, it is no longer possible to aim for (real or apparent) control by all users at all times: often another user will be carrying out a conflicting action at the same time. Who owns which pieces of text? Who has precedence? Who has permission to execute which action? Should these issues be enforced by the technology, or should the group itself work out social protocols to deal with them? Also, these issues are constrained as much by technological capability as by social preference.

ShrEdit uses a locking mechanism to protect a piece of text which is currently being edited. However, although the system informs users that a conflict has occurred, and displays the lock cursor, it does not say who the conflict is with (who temporarily owns that piece of text). Nor is there any clear indication as to what constitutes the locking range, for example, a message saying 'please move by a word, line, or paragraph'. When combined with the general lack of feedback and awareness information, the participants found this impossibly confusing and frustrating: however, their proposals were hampered by their lack of knowledge as to what was technologically determined and what could be redesigned. The problem of protecting finished work also interested the participants: both group work, and individual work which a user did not want changed. Should they have open access and honourable protocols, a strictly defined hierarchical editorship, or a locking mechanism which could be enabled on selected pieces of text? No final decision was reached, with debate ranging from total freedom to total control, as illustrated by the following snippets of conversation:

[2]This interface feature, unbeknown to the participants, is used in Aspects^{TM}, a commercially available shared editing tool.

S: Why do we have to move out of the way - isn't

of typical Macintosh™ editor 'look and feel' did create expectations of habitual Macintosh™ functionality. This led to problems when the formatting action on a piece of highlighted text differed from that expected, changing the whole document instead. This lack of consistency between other editors and ShrEdit was deemed both misleading and frustrating, not to mention pointless: S: A whole document in Dingbats - ridiculous! (Playschool log). Again, we predict that this issue will only gain in importance once the basic application has matured somewhat.

Summary of participants' redesign

The participants were most concerned with issues of awareness, communication and ownership. At all times, it was pointed out by the participants that a user must know *who* is doing *what* and *where*

Grudin, J. (1988). Why CSCW applications fail: Problems in the design and evaluation of organisational interfaces.. In *Proceedings 1988 Conference on CSCW* Toronto.

Grudin, J. (1989). The case against user interface consistency. *Communications of the ACM, 32 (10)*, 1164 – 1173.

Halverson, C. (1994). Distributed cognition as a theoretical framework for HCI: Don't throw the baby out with the bath water - the importance of the cursor in air traffic control. Technical report No. 94-03, Department of Cognitive Science, University of California, San Diego.

Harper, R., & Carter, K. (1994). Keeping people apart. *Journal of Computer Supported Cooperative Work, 2 (3)*, 199–207.

Heath, C., & Luff, P. (1991). Collaborative activity and technological design: Task coordination in london underground control rooms. In *Proceedings ECSCW '91, Amsterdam.*

Hughes, J., King, V., Rodden, T., & Andersen, H. (1994). Moving out from the control room: Ethnography in system design. In *Proceedings CSCW '94, North Carolina*, pp. 429–439.

Hutchins, E. (1990). The technology of team navigation. In Galegher, J., Kraut, R., & Egido, C. (Eds.), *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work.* Lawrence Erlbaum: New Jersey.

Hutchins, E., & Klausen, T. (1992). Distributed cognition in an airline cockpit. In Middleton, D., & Engestrom, Y. (Eds.), *Communication and Cognition at Work.* Cambridge: Cambridge University Press.

Jefferson, G. (1984). Caricature versus detail: On capturing the particulars of pronunciation in transcripts of converstational data. Tilburg papers on language and literature No. 31, University of Tilburg, the Netherlands.

Kensing, F., & Madsen, K. (1991). Generating Visions: Future workshops and metaphorical design. In Greenbaum, J., & Kyng, M. (Eds.), *Design at Work: Cooperative Design of Computer Systems.* New Jersey: Lawrence Erlbaum Associates.

Krauss, R. M., & Fussell, S. R. (1991). Constructing shared communicative environments. In Resnick, L. B., Levine, J. M., & Teasley, S. D. (Eds.), *Perspectives on Socially Shared Cognition*, pp. 172–200. American Psychologial Association.

Kuuti, K., & Arvonen, T. (1992). Identifying potential CSCW applications by means of activity theory concepts:

Orlikowski, W. J. (1992). Learning from notes: