Interactive Representations for Reflection in Group Simulations

Mike Dobson, Julie-Ann Sime & Michael Pengelly
Lancaster University
m.dobson@lancaster.ac.uk

ABSTRACT
A set of tools is presented within the theoretical framework of activity systems that is designed and developed to reify the communication acts of team members to support reflective learning. An evaluation tests the acceptability of the tools used within two multi-agent virtual reality simulations. We explore the affordance of the tools and issues around their acceptability in two experimental user populations.

Keywords
Reflective learning, sociogram, simulations, agents, simulations

REFLECTIVE TOOLS IN ACTIVITY SYSTEMS
The once contested view that learning can be effectively achieved in the reflective pursuit of operational goals (Brown, Collins, & Duguid, 1989; Schön, 1987; Wenger, 1998) is now recognised widely as an important and constructive contribution to the accumulated and available expertise of instructional design. Of course learning to perform well and developing the knowledge necessary to increase performance and transfer to novel situations in the same area of practice may not be the only desirable form of learning (Spector, Christensen, Sioutine, & McCormack, prep). There may still be other forms of instruction and educative practices that are necessary; but for the development of performative expertise there is now common agreement that learning is best achieved, has greater chance of application and is most likely to transfer when it is situated among a group already involved in the pursuit of the operational goal to which the learning is ultimately directed. When learning is situated within a community of practice the development of knowledge and of competence to perform well at an enterprise are manifestations of the active interplay between experience and ability. In Wenger's view (1998) such engagement can be most productively analyzed in terms of the dual and intimately related processes of reification and participation. In all enterprises (and he says specifically) "from ensuring our own physical survival to seeking the most lofty pleasures," learning depends on the negotiation of meaning through practice. This is often achieved in a context of highly repetitive patterns of work that can appear as only simple classification and differentiation tasks. Novelty, and therefore meaning, arises both from recognition of the similarity of patterns, maybe against a new backdrop of experience with which to assimilate them, as well as from new and unforeseen patterns of behavior.

Reifications are like good (conceptual) tools, they amplify the effect of a process while making the performance at the activity less onerous. The power of a reification lies in its succinctness, its portability, its potential physical persistence, and its focusing effect (features which also lead to its potential weaknesses). Participation and reification are complementary and can be seen to compensate for their respective limitations. Participation makes up for the possible rigidity and distance of reification. Judges, for example, are needed to interpret the law and new company policies are explained in large open meetings. Potential users may misinterpret the abstractions, seeing them as inapplicable to their own circumstances, or may be unable to re-concretize them as fluid and transferable principles. Conversely, participation alone can lead to lack of direction and misalignment and communication breakdowns and disturbances. As Wenger says, "When the informality of participation is confusingly loose when the fluidity of its implicitness impedes co-ordination, when its locality is too confining or its partiality too narrow then it is reification that comes to the rescue." In this variety of situated learning it is these two processes and their products that represent the key to unlocking enhanced meaning, development of performative knowledge, and the effective improvement of practice.

During recent years in emergency team training there have been several theoretical and practical developments in approaches to learning and training. Perhaps the two most important of these have been, (1) a shift of emphasis with regard to error from one of achieving error irradication, to one which includes strategies for mitigating the consequences of error, and (2) an increasing importance placed on the effectiveness of teams. Team training for dealing with emergencies has often emphasized minimizing error and improving team co-ordination by effective preparatory practice and drills (Byrnes & Black, 1993). Of course emergency situations occur with regularity and are
often found to be partly caused by error. A strategic approach to improving overall performance takes account and prepares for situations where errors do happen. Rather than focusing exclusively on error eradication, trainers try to develop practices which prepare learners to mitigate the consequences of possible error (Merritt & Helmreich, 1999). The instructional design shift here involves moving from environments designed for practicing pre-designed procedures, to environments where learners can experience exemplary management of emergency exigencies. Using training scenarios that demonstrate effective error management can illustrate skills desirable in emergency situations. The operatives who use this approach are expected to develop skills which help them deal with the complex and unforeseen situations which are commonly found in difficult emergency situations (Hartel, Smith, & Prince, 1991).

Often emergency situations require the co-ordination of, and communication among, decision makers with expertise in varied and complex scientific disciplines, technical logistical services as well as representatives of intricate social and commercially sensitive interests. Even at the operational level the execution of response to emergencies are likely to involve experts with different backgrounds. One response to this kind of factor is to recognize the importance that communicating situation information encourages group situation awareness and minimizes the risk of poor communication and is likely to reduce error. This shared mental model framework has been used to explain the concept of team situation awareness (Cannon-Bowers, Tannenbaum, Salas, & Volpe, 1995) and provides a rich scheme for characterizing and enhancing interactions among team members.

The team behaviors that we are trying to encourage are derived from a number of past taxonomies (Cannon-Bowers et al., 1995; Fleishman & Saccaro, 1992; Mcintyre & Salas, 1995; Nieva, Fleishman, & Reick, 1978; Swezey & Salas, 1992). We use the term ‘team competencies’ when talking about the behaviour of individuals which directly contributes to team performance. They are applicable to communications intensive environments that require shared understanding of team and individual goals, and the plans, procedures, and roles of individuals to achieve those goals. We list them here:

- mutual performance monitoring,
- giving and receiving performance feedback,
- loop closing communications, e.g., checking that information is received,
- backing-up actions of other team members with informative compensatory and anticipatory speech acts,
- adapting to resource scheduling conditions by actively balancing task-loads,
- co-ordinating action with compensatory and anticipatory communications,
- modeling team players to understand roles.

It is our claim that simulations provide an ideal environment for learning through situated practice, especially when support for reflection is also available. The claim is further reinforced in those areas where the cost of the alternative realistic preparations for situations are prohibitive. To overcome some of the known frustrations of working with open simulation environments our learners are provided guidance with carefully designed representations to help reflect on team interactions. There is no doubt that reflective thinking plays an important role in the development of self-regulatory cognitive skills. Reflection enables planning which often leads to the regulation or revision of a prospective course of activity, sometimes even before the activity has begun. We usually think of reflection as a tool for assessing outcomes after the event, often when activities have failed to lead to the desired outcome. These ideas have been in currency since Dewey (1933) and have been applied in more recent years particularly to attempts to develop informed and self monitoring approaches to scientific reasoning. Schauble and his colleagues (Schauble, Raghavan, & Glaser, 1993) developed a graphical framework and representational language for reflection in scientific reasoning. The language reifies goal and sub-goal hierarchies in an attempt to help problem solvers remain aware of reasons for their actions. It also focuses problem solvers on the relation between goals to which particular activities are directed to help users concentrate on hypothesis driven activities and not simply data driven search. Their DARN (Discovery And Reflection Notation) tool was achieved through an activity trace that was then recast into a graphical format for viewing by the user.

In a similar way we have developed a graphical trace that begins to address some of the key learning requirements for emergency executives by capturing the historical log of communication acts that occur during execution of the group simulation. These acts are the key forms of interaction between collaborating agents. Our tools provide reification of the important communicative processes, compensatory speech acts, anticipation of needs and so on, that are found to be important in the current emphasis on participation. Just as Glaser’s tools helped learners to reflect on strategies used to achieve problem solving goals by reifying aspects of the problem solving process, so our tools provide representations of the communication acts which occur among our teams. The reflective tool currently provides four representations, (1) a message list, (2) a linear representation, (3) a baton passing or floor control presentation and (4) a sociogram. Examples of the third and fourth of these appear in Figures 1 and 2. The
first is a simple list of communications which provides message start times, the operational direction (or guidance) that led to the message, the originator and recipient and the text of the message. The medium of the message, that is whether constructed by an agent or human and delivered by talk or by text is represented as a coloring of the line of fields. The data is collected as a result of direct interaction among the agents and the learners during the use of the simulators. The second representation turns the communication acts into directional lines in an elaborated and typed array of communication events. The agents involved in the scenario are presented along the horizontal axis. Vertical lines represent both the computer agents and the human participants and are colored grey. Passive agents are not intelligent but tend to play important roles in the development of the emergency. These include for example, whiteboards, diaries, flip-charts and maps; all are also included in the presentation and are shown in green. All messages and speech acts are tracked but those which are of particular importance to the development of team communication competencies are typed and annotated. Anticipatory speech acts are labeled “antic” to show the speaker anticipated the need for information before it was requested and provided it to the rest of the team. Other important speech acts are compensation and feedback. In this linear development of arrows the user can click on any of the lines to show the content of the message. This appears as a small text box along side the indicative line. The sociogram works by placing all agents (passive and intelligent) and human learners involved in a scenario in an equidistant ring arrangement. Messages collected during the scenario are filtered and shown as lines between participants. The density of communication between any two players is shown by the size of the arrow head going in any direction. From this we can see the distribution of messages relative to particular individuals in the group; a feature that can yield useful reflections on the work-load balance within a scenario. This representation allows patterns to emerge in the dialogue which relate to both beneficial and undesirable communication clusters. The sociogram also tends to illustrate the underestimated importance of shared representations (passive agents) by the use they are made of in a real situation. The sociogram representation is most effective in showing the densities of communication between roles (lots of lines joining one agent) and the linear representation is best at showing densities in time (lines closely presented in time).

These representations show the same data in different ways and therefore make best advantage of the representational affordances of each representation. Several vital instructional benefits are available from these tools for the learner to support reflective insight. The most important of these are: time and resource distribution in communication behaviors, relative significance and roles of passive and intelligent players in the construction and storage of shared mental models, and the occurrence and progress of specially typed speech acts in team communication. The relative affordances of each representation means they can be used to address certain conjectures and working hypotheses about team behavior. Such conjectures also point to aspects of behavior to look for in the analysis of interaction within a particular team.

THE STUDY

In order to uncover the user’s experience with the simulations and the reflective tools, and generate recommendations for design improvement, groups of learners from our partner test sites (an underground railway system and a nuclear power generator in northern Spain) were viewed using the tools in an observation study style evaluation. This approach has been used in learning technology evaluations for some time (Laurillard, 1987) and has been shown successful in revealing strengths and weaknesses of tools in use. Participants were selected as representative of the target user group and with sufficient distance from the development project to remain relatively unbiased. In truth, a strength of the project lay in its diversity of membership; throughout the project the end-user partners maintained a healthy skepticism toward the approach and most of their employees could have been relied on to provide critical feedback throughout. Users interaction with the simulation during the observation studies provided data which was used with the reflective learning tools. This data then gave the opportunity to provide evaluation of the reflective tools. The second evaluation effort looked at users’ perceptions of both the reflective tools and the simulations in terms of adequacy, opposition to integration, and understanding users perceptions of learning affordances. Data on these acceptability questions was collected through a co-constructive interview approach (Knight & Saunders, 1999; Lincoln & Guba, 1985) based around a ten question interview schedule.

Evaluating tools for reflective learning

At the same time that users worked with the simulations and observation data was collected, the simulations generated files of log information which was represented using the reflective learning tools described above. The capture of these log analysis files presents the opportunity to try out the reflective tools and generate some critical reaction to their use. In this sense the evaluation is akin to the qualitative accounts described by (Eisner, 1990) and should be seen as formative. The main findings are described here in terms of the function of the tools to illustrate cultural differences in communication patterns, missing communication acts, and some reflection on how the tool algorithm supports particular kinds of simulation. The distribution of communication events around the team involved in dealing with the Metro Bilbao underground railway scenario is broad. From the sociogram
representation on the left of Figure 1 we see there are no overly sized arrows between any of the agents. This illustrates that no two participants have to take an unbalanced weight of communication during the unfolding of the emergency.

![Sociograms of Underground railway simulated emergency (left) and Nuclear Power Generator simulated emergency (right). Roles include: Traffic operator, Train driver, Station Officer, Emergency Centre Director. In the Nuclear Power installation simulated emergency (right) the roles include: Shift Chief, Emergency director. Head of logistics, Technical Evaluator, Head of Health and Safety, and Local Government Coordination Centre.](image)

Initial analysis of the representational affordances of these tools suggest two forms of pattern matching benefits, (1) comparing actual team behavior with desirable team performance, and (2) the comparison of social and cultural forms of interaction in the execution of the emergency. The diagrams provide a very quick and simple method to check that all the proper communications which should take place during an emergency actually happened. In our underground simulation we can check back against the emergency protocol in (Maseda, Arenillas, Garzon, Arcos, & Castano, 1999) where we find that the train driver is supposed at one time to communicate with the station supervisor. This communication is clearly not present on the sociogram and shows how these tools can be used in simple error checking with a team in post-hoc analysis. In comparing the nuclear demonstrator with the underground railway demonstrator we can see immediately from the sociograms (Figure 1) that the role of the emergency director is one principally of informing others around him and with relatively little feedback from those other role players. Compared with other role players he receives a similar amount of information but his role in sending this information to others is perhaps six fold that of any other agent. From looking at the baton passing representation from the nuclear demonstrator shown in Figure 2 (lower) we can see the extent to which the baton remains in the hand of the emergency director. The horizontal lines drawn along the emergency director axis of this graph (shown as DE) show situations when the emergency director communicates and then the next communication to follow was also from the director. The director did not have to wait for further information to continue. The spikes on this diagram show that the director mostly generates the questions for new information from his team mates and is quickly provided with it (shown as a quick return spike) whereupon control and stability returns to the director. The two scenarios display very different communications patterns, aspects of which can be seen through each of the three representations provided by the reflective learning tools.
We have also begun to see some limitations of this early prototype representational scheme. There are many role players in the scenario who do not appear on these diagrams and there are differences between those that appear on the sociogram and those which appear on the floor control diagrams (see below). The reasons for this are worth explaining. The first constraint is determined by the simulation designer. Our simulations have so far been developed to illustrate the training of an individual or specific combination of learners. We have rather assumed that learners benefit most from learning about roles with whom they have direct communication during the emergency. Roles which are not in direct communication with the trainee are not implemented in the scenarios and cannot appear in the sociograms. For example, the communication between the SOS agent and the police, the ambulances, the hospitals the red-cross and the government mediation services, even the communication between the driver and the public who are the first to alert the driver to the emergency, are not shown on the graphic. This is because the designer considered second order roles as non-functional to the individual needs of the target individual in the training scenario. On reflection this choice was seen to have disadvantages to the learner. It was discussion around the reifications that led to this outcome. Designers need to limit their environments to avoid extraneous programming effort and to reduce overhead to the learner. However our current practice to limit implementation to roles with direct contact with the trainee does not suggest a principled guideline and will need to be improved on in
the future. In the floor control representation shown in Figure 2 communications which do not generate immediate
responses from anyone do not appear. A lot of the communications in teams is informative, broadcast for
information to all, and therefore unlikely to generate immediate responses, and will not appear on the floor control
diagram in its current implementation. There is also currently a limitation in the algorithm for presenting the data.
At present, each time an agent speaks the time is noted and a line drawn back to the intersection of time and
identity of the previous speaker. In this way the continuity of speaker contribution is maintained through the team
dialog. This representation is based on use within a closed group of speakers, often in the context of face-to-face
management meetings, or board meetings. The representation is used as a tool for analysis of power relationships in
groups. However in our simulations agents often talk to other agents who do not speak to other agents within
scenario. For example, when the SOS agent receives a message from the emergency center supervisor in the
underground railway simulation, this is likely to precipitate a message to the ambulance or the government
mediation service. But since they are not shown on our demonstrator, the message from the SOS agent will not be
shown, and the relative importance of his role will not be seen. The baton passing representation is therefore
contextual to the scenario and perhaps more useful to the scenario trainee than to other team members.

Investigating Acceptability
However usable and effective the tools and demonstrators may be, the value of the innovation will be greatly
reduced if user’s concerns and working practices are not addressed. Several acceptance issues surfaced during
development and while planning the evaluation. The acceptability interviews elicited information of this kind from
strategically chosen data sources concentrating on a small selection of individuals representing key stake-holders in
the organizations, interviewed with a pre-planned interview schedule but open-ended and interactive approach.
The selection of interview respondents was influenced by the sensitivity of the work culture and the potential for
respondents to provide useful information. The simulations may seem to present few challenges of “acceptance” for
the user group since the prospective users in each partner organization are active participants in the development
project. Certainly within the partner organizations the intended demonstrator user group consists of senior
managerial staff who were committed to the project. Acceptance problems are more likely to be seen in reactions
from other branches of the same organization and public safety organizations concerned by the effectiveness of the
distributed simulations. The general public and the media may have reactions to the demonstrator implementation
that could eventually turn into problems for technology exploitation. As the profile of emergency response training
increases in the public sector, the media and perhaps other branches of the partner organizations are likely to take a
more direct interest. While our senior manager participants have benefited from the project throughout, they often
maintained a strong sense of their operational goals and were at times openly skeptical about the approach. As
interviewees this group represent a self interested, responsible and critical group with very high levels of knowledge
in their target domains. The questions used to lead this interview were as follows.

• How do the tools address learning requirements and needs experienced in your workplace?
• What are the main limits of performance or concerns about training in your workplace and how are they
  addressed by the tools?
• What aspects of the tools and simulations are most innovative and would cause most change to your
current practice?
• Does the manner in which the simulations and tools are designed suggest they will be understood by your
  co-workers?
• Is reflective practice an approach to training which you feel you understand and is appropriate to apply to
  training for emergencies?
• Is it important for co-workers to develop shared mental models of emergency situations and emergency
  situations, and why?
• Does the training approach make it difficult to use results to address the compliance of safety standards
  which are legislated by your employers?
• Would the approach taken in the demonstrators be welcomed by other branches and offices of your
  organization?
• Will managers, co-workers and competitive organizations find the approach difficult to understand?
• Will public perceptions of your approach to dealing with safety matters of training be enhanced by learning
  about the approach taken in the simulations?
• Do you believe that public perceptions in such matters are easily influenced?
Three groups of issues were considered likely to affect the acceptance (or rejection) of the technology; (1) the industry need, the niche provided by the technology and the threat of the technology to current practice, (2) users’ level of understanding and belief in the instructional approach, and (3) the organizational and social climate toward the innovation. The interview schedule included three questions arranged in each of these groups. A final question provides an opportunity for the informant to add their own themes to the discussion.

The results of these interviews provide valuable insight into how our user communities viewed these issues as well as bringing direct information to bear on the specific questions. In our view the most valuable results lie in the user’s unexpected interpretations of the questions. It is in these responses that the acceptance and integration challenges are likely to be found. In the first and second questions we asked how the simulations addressed learning requirements in the informant’s workplace. From the questions we hoped to find out how potential users viewed the simulations as an advantage in meeting their cognitive and performative needs. To describe this match we expected respondents would describe their needs, describe what they thought of as the advantages of the shared agent simulations, and then show how the needs and affordances co-ordinated. The second question specifically asks about the needs for performance improvement in the organizations. In fact the interviewees were hesitant to discuss needs or affordances in any specific way and preferred instead to provide information about the satisfaction of assumed and unsaid needs. Users, for example, commented favorably about the facility of the simulations to present life-like and at times unexpected situations. They thought it attractive that the training could be run without the need to spend a lot of time assembling operational workers. Users appeared to like the idea that “a deeper review of practice (could be done) though the reflective practice tools” and that the simulations allow “multiple training and multiple roles.”

What at first glance is surprising is that none of the respondents were forthcoming about how the specific local performance or cognitive needs were creatively addressed by the technology. The sensitivities of safety conscious industries, in particularly of the nuclear power generators, are at work in these responses.

The simulations use some relatively unusual instructional strategies that are unlikely to be common among the user populations. We identified several aspects of the design that we thought could cause concerned reaction from the user populations. For example, the requirement for learner’s to take more responsibility over their performance improvement through reflective practice, and the recommendation to train for other workers roles, seemed to us as the most likely threats to successful integration. We wanted to know if the user populations saw these differences in expected learner behavior as problematic and how they believed any problems could be resolved for integration. Again, the surprising reaction was that our pedagogical innovations were much less salient for users than were the generic technologies; the virtual reality, the agents and communication with human learners. It seems interesting to us that while we are advocating a training system without any means of criterion referenced testing and placing most of the emphasis on learners to use the reflective practice tools to manage their progress, that the aspects of the simulations users thought most likely to change current practice in training were the representations of shared practice.

The next set of questions turned to the manner in which the demonstrators are designed and the extent to which workers near to the users in their workplace would understand the role and function of the training. We wanted to identify potential sticking points with other groups in the target organizations. The interviewers asked about how the instructional methods used in the simulations might affect or be responded to by other areas of their organizations. Responses here fell into two categories that are again most interesting for their unexpected interpretation of the question. Initially, one group responded that the demonstrators could benefit from more support either with the user manuals of with their overall usability, and the second group projected surprise at the question since they thought the technical aspects of the systems are not difficult to understand. Far from seeing the potential of the innovation to challenge current performance practices, approaches to validation and role descriptions, the respondents understood the possible barriers to integration to lie entirely in the systems’ ease of use.

Later, the interview pursued the specific affect of the reflective practice approach on work practices and training. Responses ranged from characterizing the approach as “totally subjective” through suggesting it be combined with objective assessment, to recognition that it presented the most opportune time to consider performance after stressed performance. Others thought that while it might be useful were concerned that “nothing should come between fast and effective response to the emergency,” implying (presumably) that reflective learning may slow the response process and inhibit performance. The other key pedagogic aspect of the demonstrator design lies in the support for the development of users’ shared mental models. Contributions to the interview here, showed that the technique was understood and welcomed, respondents claimed the technique is the demonstrator’s “most unique feature”, and “without it the traditional training would be just the same,” that it is “good for illustrating the complex co-ordination skills of the team leader” and that “since peer staff change regularly, it helps to build continuity and confidence across the teams.” After pursuing the relation between pedagogic approach and organizational integration beyond the initial questions the user group were able to offer informed and convicted opinions about their value. The connection between the two was far from salient among the user group. The third set of issues dealt with higher
level organizational issues; compliance, legislative criteria and with regulatory bodies. The socio-cultural view on learning and performance clearly views learning not as an effect of training but as a reaction to it (Wenger, 1998). We asked about the implications for acceptability arising from government processes and legal requirements in the industries to determine the organizational expectations for managing performance improvements. Responses identified the need for objective learner assessment to illustrate results to regulatory bodies and to illustrate the benefit of the tools and approaches to improving performance. Unfortunately, as above, confidentiality issues which manifest dramatically in the nuclear industry again intervened in the discussion about regulators and their emphasis on performance, training and compliance. The emphasis on objective measurement indicates a tension between the organizational expectations and the approaches used in our simulations. For example the attempt to create shared mental models in team members, combined with reflective practice, yields no obvious performance measurement techniques, but the approach is specifically designed to help with aspects of operation in emergencies where performance is recognized in the research to be poorest (dealing with complex, non-recurrent and unforeseen situations).

The interview progressed to discuss acceptance reactions to innovation from groups quite remote from the simulation’s central operational activities. Interviewees identified other analogous branches of the industries which might find the technology useful but suggested that their acceptance would be dependent on further development in system usability. Again the results here show how difficult it is to imagine the affects of implementing such technologies into workplaces even for developers and full time operatives. Both the target domains are closely connected with government organization and are no doubt connected in very many ways, through practice, funding, representation, supply and other mechanisms to other branches. Respondents seemed to find it very hard to imagine any of these links and perhaps even more strange to be asked about their public constituencies. Responses to this topic included, “the public always like new technologies and new techniques and project participation is considered positive.” One said “the public will be pleased by the system that they don’t have to be disturbed by the drill and practice of the normal emergency training.” The main category of public constituency reported were the industry customers (e.g. train users and electricity consumers). It is interesting that emergency workers appeared less concerned by public confidence in the practice, performance and management of emergencies, and more by thoughts of convenient train use and reliable electricity supplies. After all, in both our target domains the responsibility for safety procedures is directly in the interests of the general public, not the train users or the electricity purchasers, but the general public too. When asked about public confidence, operatives replied that practitioners were “not qualified to answer the question” or that the issue was “outside the scope of our work” and that “while technical effort expended by the developers was likely to be appreciated by the public, their general perceptions were more likely influenced by political and social criteria.” Public perceptions, their relation to changes in practice at the application sites, and the public’s response to change was not regarded as important to the safety operatives or the instructional trainers in the respondent groups. Respondents certainly felt these questions were not in their domain, but since public confidence and safety performance are a function of training approach we suspect the problem may lie in communication between public relations and safety training sections of the organizations. The results of the acceptability study are complex and responses appear often less interesting in first analysis than they are in interpretation of the questions. By way of summarizing what the evaluators believe are the most useful outcomes we have provided a list of markers as follows.

- Effects of instructional approaches in changing organizational practices were not obvious to operatives or developers without strong prompting.
- The technical design; virtual reality and agent modeling, were perceived as more innovative and novel than the instructional approaches.
- Other power generating sectors should consider adopting the technology
- Cross-training to achieve shared models is a useful strategy
- The shared mental models are the most unique aspect of pedagogy.
- Operation of safety critical practice is governed by legislative authorities which makes compliance as important as performance.
- Safety organizations may experience sensitivity toward learning needs analysis, making performance data difficult to access, also hindering studies of learning effectiveness.
- Informants perceived their public constituencies as customers with little interest in safety management or training.

In summary of the acceptability study, we set out to understand the factors involved in the potential success of integrating similar reflective practice tools in organizations. We attempted to understand their training needs, identified performance gaps, the problems that novelty in training would bring to practice (specifically of shared
mental models and reflective practice), what effects government regulation might have on training, performance and needs identification, and how far acceptance could be visualized and considered. The outcomes reflect a complex organizational and theoretical context for integration and particularly illustrate that operatives can be unaware of the wider activity systems in which they play an active role.

CONCLUSIONS
The domains in which we work are representative of tremendously important emergency and safety issues which are embedded within complex organizational contexts. We know that distributed human and machine systems are susceptible to error and failure and that emergency situations are often unpredictable leaving operators with no well practiced behavior and without contingency for dealing with the consequences of error. The instructional design guidance for designing learning environments for complex domains (Merrienboer, 1997; Spector & Davidsen, 1998) provides guidance in the matter of performance and complexity of task oriented training but little in terms of team performance and specialized approaches to the emergency field or to the organizational factors that influence successful adoption and integration. The design of our reflective tools are rooted in several key ideas. The approach assumes that performance results are effectively enhanced during the pursuit of operational goals and that simulation environments provide realistic and valuable context for performance. Within the simulations we subscribe to the view that reifications and shared representations of practice help in communicating and reflecting on practice. Furthermore we claim that preparing for emergencies (and for all non-recurrent skills) must include not only whole task performance but also preparation to mitigate consequences of error, and that training teams to exhibit competencies in team skills can be such a mitigating strategy. We then described the design of tools for reflective learning based on several graphical formalisms which together are designed to be used in coaching and post-session debriefings.

The evaluation consisted of two studies. In the first, two groups of users worked with the simulations while data on use was recorded by eyewitness evaluators from the target user communities. User interactions with the simulations were stored to a logging system from which the data can be represented with our reflective learning tools. The second study gave opportunity to consider user’s perception of adequacy, opposition to use and integration and potential acceptance of the simulations. The reflective learning tools showed that communication patterns varied widely between the two simulators illustrating differences both in practice, culture and process in the two scenarios. Organizational factors in communication during unfolding of emergencies are made visible in these illustrations that may not be obvious without these carefully constructed representations. The tools have potential to help users understand the entire sequence of an emergency, not just their own role, and this should promote the development of shared mental models and better performance in teams. Through implementing the demonstrators and the tools for reflective practice we have introduced to the community of emergency preparation specialists a series of shared interactive representations and reifications for enhancing reflective performance that are based on contemporary and tested theory in the field. Eventually these tools and their derivatives will be of widespread benefit to the target community providing a cost-effective development solution to similar domains. But as a consequence of this ambitious and itself reflective project we have generated the start of a set of our own questions in the field which we think will be usefully executed in the coming years. We believe the key questions to emerge from the project involve the affordance of the technology and the instructional approaches. We believe it would be of direct benefit to our exploitation partners for example to know, (1) what factors in team constitution make dealing with unexpected non-recurrent situations more effective, (2) in what ways analysis of real emergency situations could contribute differently to traditional drill programs in the creation of useful simulation environments, (3) for which types of teams and industries are the shared mental model training approaches likely to be most helpful and (4) what information design factors in representations are likely to make sharing more beneficial to the community of practice.

ACKNOWLEDGMENTS
We would like to thank Brian Slator of NDSU Fargo for comments and the anonymous reviewers at USAF for their comments on a related paper. The contribution of all partners in the ETOILE project is acknowledged but the views expressed within are not necessarily agreed by all partners. This work was partly funded by the European Commission under the Fourth Framework Program (ESPRIT: Project 29086 CEC DGIII Information Society). Programming of the representation tools was mainly done by Walter Onyino in Computer Science, Lancaster University.
REFERENCES


