COLDEX

Collaborative Learning and Distributed Experimentation

Information Society Technologies Programme
Project number: IST-2001-32327

Evaluation Plan I: Methodology and Examples

Deliverable Number:	D2.3.2
Contractual Date of Delivery:	M18
Actual Date of Delivery:	M30
Version:	Final
Work-Package:	2
Lead Partner:	VXU
Contributing partners:	INESC-ID, UDUI
Authors:	Björn, Marianne; Karlsson, Marine; Milrad, Marcelo; Ulrich, Hoppe; Wichmann, Astrid & Otero, Nuno
Contact:	marcelo.milrad@msi.vxu.se

Table of Contents

1 Introduction

This paper is the final deliverable from workpackage 2, based on task 2.3: "Definition of an evaluation plan considering the learning activity design". It contains and describes several evaluation methods and gives examples of how these can be applied. A more specific evaluation and test plan will be elaborated as part of workpackage 8, (deliverable D.8.1.1). This future deliverable will provide concrete examples and recommendations on how the different partners in the COLDEX project can plan and realize the evaluation of their activities.

The COLDEX project aims at designing innovative learning environments in order to support a wide range of global open learning activities within the scientific domains of astronomy, biodiversity, chemistry and seismology. This goal is achieved by combining innovative pedagogical approaches (e.g. Challenge Based Learning, CBL) together with the support of a variety of modelling tools and experimental scenarios.

Interactive tools to create and augment collaboration enable the plain re-use of objects by an integrated learning object repository (LOR) that represents the "group memory" of the community. This repository supports a contextual search adapted to the necessities of each community. The different tools and learning objects can be accessed through specific digital experimental toolkits and through the COLDEX on-line portal.

The specialty of the COLDEX project in CSCL terms is that it aims at developing methodologies to create and maintain large learner communities around complex experiential phenomena rather than focusing on small highly controlled laboratory situations. Local learning communities will exchange their ideas and work in an "Open User Scheme" (OUS). The local learning communities share a rich everyday context. Teachers and students are expected to take an active role in creating, filling with content and structuring this network.

The following sections will provide an overview with regard to those aspects that should be considered while evaluating a COLDEX related learning activity. Section three introduces the ideas behind the use of activity theory as a framework for evaluation. Sections four and five present and discuss which components of the educational and technological aspects are important to be assessed. Section six describes a number of techniques for data collection and analyzing methods. Section seven concludes this deliverable by illustrating how some of the ideas presented in this document can be used for a particular educational scenario. The ideas presented in this deliverable are described having in mind a "meta level" perspective. The main reason behind this view is for providing the epistemological foundation of the evaluation process. The next deliverable will give a more detailed description on how the evaluation plan should be conducted.

2 Basic aspects about evaluation

The COLDEX evaluation will be done in different countries and by several different persons with different perspectives, backgrounds and goals. Therefore, it is essential that evaluators at the different COLDEX sites discuss and elaborate the evaluation activities from an overall project level, and from the project partners' own local perspective.

Nash, Plugge and Eurelings (2001) discuss the complexities involved in preparing, managing and evaluating educational technology initiatives. They state that there are four prevailing misconceptions concerning projects in this field, namely that there should be (1) a common goal, (2) a common problem, (3) a reductionism's simplicity and (4) that traditional research paradigms should be useful. The misconceptions are, according to these researchers, due to the following reasons:

(1) *A common goal:* This is a misconception because a project has different stakeholders (i.e. "people who care about or have something at stake in how the problem is resolved") that sometimes are working in different domains, maybe with conflicting or hidden goals.
(2) *A common problem:* This is a misconception because since project participants deal with different problems depending on the differences in their specific domains it is likely that they work from different perspectives.
(3) A *reductionism's simplicity:* This is a misconception since it is not possible to reduce the project's complexity in such a way that it becomes a simple problem. The authors claim that it is common that a complex problem is dealt with as a "tame problem" to evade the complexity and that it therefore becomes subject to traditional, linear, problem-solving methods that the project participator is more familiar with.
(4) *The usefulness of traditional research paradigms:* According to Nash et al., there is a growing opinion that traditional evaluation methods with randomized control groups to establish counterfactuals should not be used for evaluating complex multifaceted activities.

Nash et al., (2001) claim that traditional research paradigms may depend on the underlying approach of the researcher. According to Jerkedal (1999), four major dimensions can be identified where different research paradigms have their own conception on what is right or wrong (figure 1):

- *Positivistic or natural science vs. hermeneutic or ethnographic*
- *Affected by values vs. devoid of value judgement*
- *Pragmatic, highest possible use vs. scientific, highest possible precision*
- *Pluralistic, democratic vs. controlled from the management*

These different views are illustrated in figure 1. The middle box represents a possible intermix between two opposites. Jerkedal (1999) stresses the point that the dimensions do not request an either-or application but can many times contain parts from both opposite poles. The evaluation questions together with the research paradigm hold by the researcher, will of course affect the lens through which the researcher is looking.

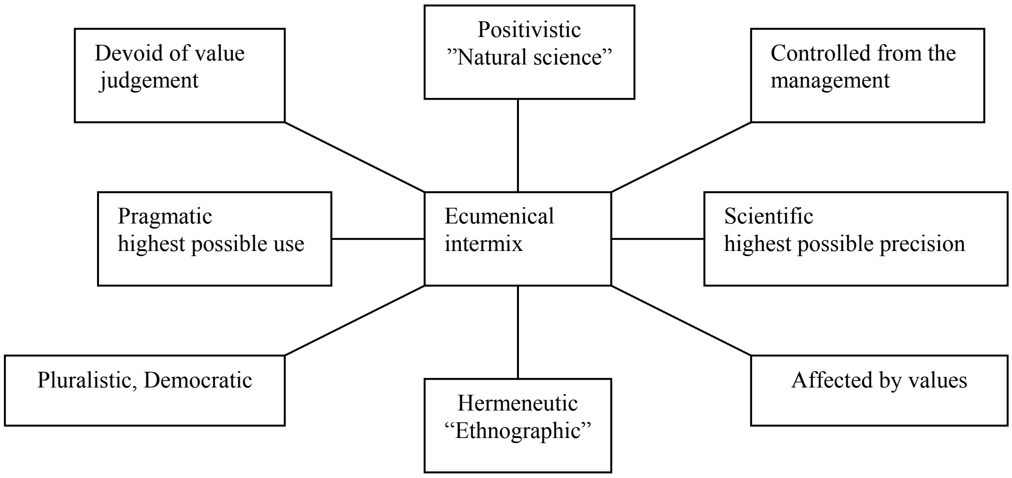

Figure 1. Different views on evaluation, from Jerkedal (1999).

The view on learning and teaching encouraged by COLDEX is based on social constructivist and sociocultural assertions, and stresses real world phenomena, with authentic problems in authentic contexts. Evaluation methods used in controlled laboratory experiments mainly applied in behaviorist and cognitive psychology sciences are therefore not in line with these ideas.

Koschmann (1996) claims that the emerging paradigm in instructional technology is Computer-Supported Collaborative Learning (CSCL). CSCL ideas are based on socially oriented theories of learning that involves cultural context, language and other social issues. According to Koschmann, CSCL research tends to be descriptive and to focus on the participants' collaboration as expressed through their words during the learning process and in their account of the work, and on the artefacts used by them. Thus, we suggest that the evaluation of the COLDEX activities will be consistent with these ideas. In the following section we give an overview of Activity Theory and related ideas that can be used as the foundations for the evaluation process.

3 Evaluation framework

COLDEX advocates the use of Activity Theory (AT) as a theoretical framework for the evaluation of the collaborative settings and educational activities (learning processes) of the project. AT may be complemented with methods, e.g. from the area of ethnomethodology and sociology that can be used for collecting, processing and analyzing data. Examples of these methods are ethnography and conversation analysis (Preece, Rogers and Sharp, 2002; Schuler and Namioka, 1993).

One of the major objectives of the sociocultural approach is to elucidate the relationships between human action and the context with its historical, cultural and institutional influence (Wertsch, del Rio & Alvarez, 1995), that is to analyze human beings in their natural environment (Kaptelinin, 1996). These influences are important to examine further in connection to the inter-cultural exchange in the COLDEX's network.

The focus of AT is on one activity system or on networks of activity systems and the processes and the tensions within and/or between them. The contextualized activity system is commonly presented as a triangle (see figure 2) with the nodes: subject (point of view for the analysis); object (acted upon); outcome (transformed into), as mediated by the nodes: community of practice; rules; division of labour; and tools or artefacts (Engeström, 1987). An example of an activity system could for example be a local student group studied on an intersubjective level. The group activity, for instance for the VXU scenario, can be expressed in the following way:

1. *The subject* – some students from a local class
2. *The object* – learn about growing plants in space
3. *The tools* – the bioTube, BioBlast, CoolMode, and others
4. *The community of practice* – the rest of the class, the teacher, (staff from the Science Center)
5. *Division of labor or roles* – project leader, secretary, collector of data, etc.
6. *Rules* – curriculum objectives, teacher's direction, Science center's open hours, etc.
7. *Outcome* – real plants, a report and/or learning objects for LOR

Though it is the nodes or entities that you define, it is important to remember that the connections between them are as interesting as the entities themselves. An activity system is dynamic and complex and can contain constraints in many places. Examples within the system could be that the subject has conflicting ideas about the object or expected outcome. Constraints can also come from outside: from other activity systems like student groups from the same or other classes; from the pedagogues or from other schools.

Activity theory sees activities as dynamic and changing and differentiates between processes at three different levels depending on their orientation (figure 3), namely: activities; actions; and operations. They are defined as; activities if they are oriented towards motives, actions if they are oriented towards goals, and operations if they are determined by the actual conditions of an activity (Kaptelinin, 1996).

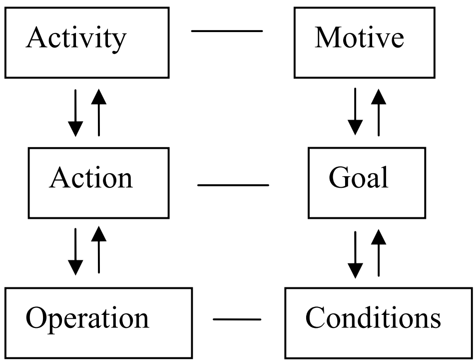

Figure 3. Levels of an activity, (Kuuti, 1996).

In other words, activities are processed during a longer time and are realized as individual and cooperative actions, which consist of chains of operations. The operations are answers to conditions faced during the action (Kuuti, 1996). Leontev states that an activity is realized through action, and that an action cannot be understood without the frame of reference or context provided by the larger activity system (Barab et al., 2002).

Rogoff (1995) suggests using three levels when studying human actions: intrasubjective level or individual; intersubjective level in e.g. a student group; and institutional or context level. Preferably, the focus is on one of the levels but with the other two kept in mind and considered. This corresponds with COLDEX approach to use AT to model and to understand the complex relations between: individuals in different roles; groups at different levels; and the "social inter-operability" of tools provided (see COLDEX project description). There will of course be focus on different levels at different times depending on the purpose, the learning activity and the learning environment. Here each research group must decide where to put the main focus. Within COLDEX the following activity systems could be interesting to analyse:

- *Individual student's learning activities in a local/global community*
- *The student group's learning activities in a local/global community*
- *The teacher/s' practise in the local/global community*
- *A science centre's or observatory's participation in a particular scenario*
- *Researchers' dissemination of research results*

It would also be interesting to compare one activity system over time, e.g. teacher's practice to see if the practice changes over time or through participating in COLDEX activities.

An evaluation of a learning activity using AT should consider the whole complex of the educational activity. Depending on the access to resources each evaluator has to narrow down to the level that covers evaluation questions relevant to the researchers and the research areas expressed in the project. The choice of method for evaluation can for example depend on who will do the evaluation. We identify the following evaluators:

- *The researchers from the different COLDEX partners.* Researchers will have different goals and different motivations for doing an evaluation. The text below suggests some possible interesting subjects and evaluation methods.
- *The teachers involved in the learning activity.* Teachers will be involved in the design of the learning activity and are most probably interested in evaluating the students' use of the environment, pedagogical implication of the CBL challenges and integration of activities into the curriculum.
- *The students involved in the learning activity.* The students will most commonly be doing self-evaluation during and after a CBL activity. They will maybe also do peer-to-peer evaluation.

There are mainly two perspectives that are interesting when evaluating a COLDEX Learning Activity Scenario, namely technological and educational. These two areas are of course intertwined since technology supports learning and learning is affected by technology. Knowledge is socially constructed, it is therefore also important to remember that the context is equally important.

A particular concern is that different cultures may have varying attitudes toward collaboration (Hofstede, 2001), and negotiating these differences may be difficult in distributed, technologically-mediated contexts. An analysis of a learning activity should therefore consider the influence that culture and history have.

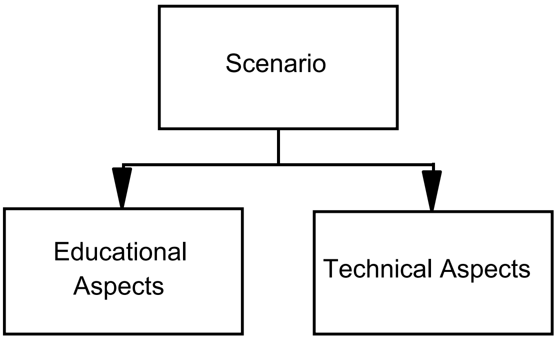

Figure 4. The components of a scenario, from the deliverable Learning Activity Design (D.2.3.1).

The ontology used in the document Learning Activity Design (D.2.3.1) specifies how a learning scenario can distinguish between Educational and Technological aspects (figure 4). In the following sections we will discuss these issues and their relation to evaluation aspects.

4 Educational aspects

The Educational aspects include Activity, Community, and Type of Learning. The Activity consists of an experiment model with educational aspects, an organizational model (with collaboration mode, communication type, task structure and mode of presence), and an outcome. The Community can be local or global and may already exist or may emerge during the Activity. The Type of Learning includes the objective, the definition of a challenge or problem and the space of learning. The features of a learning activity, as described in the table below, serve as possible foci within the COLDEX evaluation.

4.1.1 Activity

CSCL focuses on group activities, the importance of artifacts, and that the artifacts such as spoken, written and published texts capture newly constructed knowledge. Stahl (2003) suggests four foundational issues from CSCL to focus on, namely: *(1) collaborative knowledge building; (2) group and personal perspectives; (3) mediation by artifacts; and (4) interaction analysis.*

Bereiter (2002) identifies progressive discourse and investigation of ideas as “conceptual artifacts” as central for knowledge construction. Chan and van AAlst (2003) point out that the knowledge building perspective stresses that knowledge is a social product and that it involves the practice of knowledge elaboration, creation and advancement.

Scardamalia (2002) outlined 12 characteristics of knowledge building, which can be used to evaluate if a group is actually engaged in this process (Law & Wong, 2003). These features are described below:

1. *Authentic problems*
2. *Improvable ideas*
3. *Idea diversity*
4. *Rise above*
5. *Epistemic agency*
6. *Community knowledge, collective responsibility*
7. *Democratizing knowledge*
8. *Symmetric knowledge advancement*
9. *Pervasive knowledge building*
10. *Constructive use of authoritative sources*
11. *Knowledge building discourse*
12. *Embedded and transformative assessment*

Van Aalst et al., (2002, in Law and Wong 2003) assessed students’ development by using the following criteria: if they were working at the cutting edge; progressive problem solving; collaborative effort; and identifying high points. Law and Wong (2003) suggest the use of developmental trajectories to trace the learner’s discourse while working collaboratively in knowledge building.

4.1.2 Community

As Littleton and Häkkinen (1999) emphasize, much of the work on collaborative learning has focused on the collaboration within groups. In order to understand these processes we need to recognize the powerful influence of the context. From this perspective, COLDEX aspects related to different languages and cultures deserve special consideration. The local communities will exchange ideas and results in a network established by the “Open User Scheme” (OUS) and global learning communities will be established that have to build up a common context and a common culture.

There are several levels of collaboration in COLDEX. The global collaboration will develop through the OUS dissemination activity. The project starts with local communities which share a rich everyday context as e.g. in a school class or in a study group of academic students. These basic groups are collaborating and using networked interaction. The local communities will exchange their ideas, results and problems in the international network, established by the OUS.

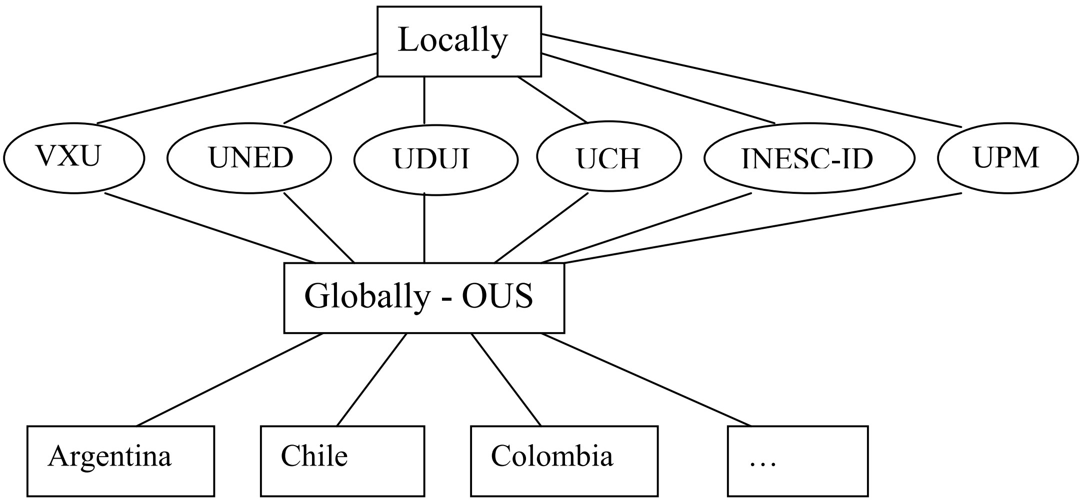

Figure 5. The organization of the COLDEX network.

The creation of global learning communities is a central issue for COLDEX. The idea is to build communities bottom-up, starting with the local communities and let these establish contacts with other communities in different countries (figure 5). The evaluation of the collaboration within the OUS should be done locally and globally. There is still a need for COLDEX to define common criteria to study for the sharing and collaboration through the global network.

Discourse analysis is a tool that can show the interaction between community members and how they co-construct intersubjectivity. Conversation analysis is a method used in order to analyze students’ conceptual change (Roschelle, 1996). Learning trajectories may be used to follow the different levels of collaboration over time in integrated learning activities (figure 6). Collaboration could be in a large group, e.g. a school class that receives from the teacher a challenge that includes individual activities, divided into smaller groups they work collaboratively with the challenge. The community could be all students and teachers that are involved one way or another in a project at e.g. a Science center.

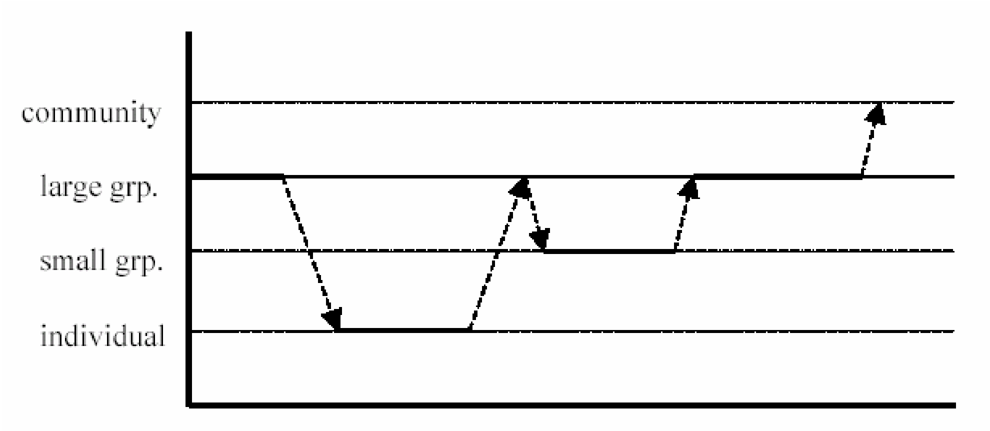

Figure 6. A learning trajectory showing levels of collaboration over time.

The example episode does not comprise a concrete learning activity on the community level, yet in the end it produces learning objects, which are exported to the community archive, the LOR. The archive may be accessed on and between all levels. Session handlers together with tools will generate and maintain metadata, which can be used to automatically index learning objects when they are put into an archive.

Christiansen (1996) points out that you may observe and interview participators in a community but you will not understand why they use artifacts the way they do until you understand what activities that are involved in their practice. She worked with a field study and made an activity topology based on data from observations and interviews in order to be able to interpret the data and identify potential conflict of views. The typology was made by asking what kind of artifact would mediate which activity, this to overcome analytical problems of interpreting the different activities within the community of practice.

The structure of a networked learning environment can be investigated using a method called social network analysis (Nurmela et al., 2003). The method can show activity sequences, interaction patterns and roles. The analysis is made on automatically collected logs with information about the user, the time, the location in the environment and the action. Nurmela et al., (2003) studied a social network in terms of density and centralization. Density describes the general level of cohesion and centralization to which extent this cohesion depends on particular actors.

4.1.4 Type of learning

The learning workflow associated with COLDEX includes activities where the learners are expected to identify/generate problems to be solved. Students then will learn, develop and apply relevant knowledge and skills through progressive problem generation, framing and solving. The different learning activities require learners to identify research questions and variables, set hypotheses, build and construct experiments, test results, analyze observations and then refine hypotheses and casual variables accordingly. All these activities correspond with many processes associated to scientific inquire learning.

The analysis of scientific inquiry learning activities used by Germann, Haskins, and Auls (1996) suggest to use the following generic categories in order to evaluate scientific inquiry tasks:

- *Formulating a hypothesis*
- *Designing experiments*
- *Controlling variables*
- *Providing evidence*

Chinn and Malhotra (2002) claim that these categories could also refer to simple inquiry tasks, however they presume a certain level of autonomy within the student activities. The independent design and conduction of an experiment design, involved a higher degree on self-regulation than is usually required of students in school tasks. The evaluation criteria developed by Chinn and Malhotra (2002) and Germann et al., (1996) analyze a particular environment by determining which student activities are supported within the tasks developed. In order to characterize student activities on a more conceptual level, cognitive processes can be analyzed. Sins, Savelsbergh and Joolingen (2003) used a coding scheme adapted from Stratford to analyze reasoning processes that are required to solve certain modeling tasks. These reasoning processes involved the following cognitive activities:

- *Evaluate*
- *Explain*
- *Quantify /specify*
- *Reason*
- *Analyze / identify*
- *Read /paraphrase*

For the analysis of cognitive processes within scientific experimentation tasks Bloom's taxonomy (1956) and thinking processes developed by Ossimitz (2000) seem to be useful as a basis for evaluation to focus on thinking processes. Blooms's taxonomy includes six levels of thinking: knowledge of terminology; comprehension of meaning; application; analysis; synthesis (of prior knowledge and skills to produce something new); and evaluation or judging. These levels are the basis for higher order thinking skills.

Roschelle (1996) proposes a process in order to analyze students' conceptual change while collaborating and constructing shared understanding. The process includes (a) the construction of a situation at an intermediate level of abstraction (b) interplay of metaphors in relation to the constructed situation and each other (c) an iterative cycle and (d) the application of progressively higher standards of evidence for convergence.

Research suggests that learning could be more productive if the students reflect on the metacognitive aspects of their own learning and thinking (de Jong, Veldhuis-Diermanse and Lutgens, 2002). Anderson (2002) has proposed five main components that give evidence for metacognition in learning situation, namely:

1. *Preparing and planning for learning*
2. *Selecting and deciding when to use a learning strategy*
3. *Knowing how to monitor strategy use*
4. *Knowing how to orchestrate various strategies*
5. *Evaluating strategy use and learning*

5 Technological aspects

In this section we give an overview of those issues related to the evaluation of COLDEX's technological aspects. COLDEX's tools can be divided into two distinct components: Workspace and Resources. A COLDEX workspace consists of one private part and one shared space. The content from the workspaces are saved in COLDEX's learning object repository (LOR). Resources include: software tools, DExTs; and other material, e.g. telescope, sensors, CoolMode, Lego Mindstorm and Active Document. DExT can be seen as a package of resources which includes:

- *Experimental instructions,*
- *Scientific background*
- *Modelling and simulation tools*
- *Access to real data*
- *Visualization and collaboration tools*
- *Initial challenges*

The current evaluation exercise presented in this section comprises two distinct levels of analysis since it is not only important to evaluate the technological parts and the tools *per se*, but it is also essential to look at the system as a whole. Thus, this section is divided into two parts. On the first part, a macro level perspective is presented. The aim here is to provide a framework to help look at the whole system and see to what extend do the different parts of COLDEX meet the learning requirements established initially (see for example, the sections above or, in detail Deliverable 2.2.1). On the second part we discuss evaluation methods and strategies from micro-level of analysis perspective, where tools and representations are the focus.

5.2.1 The wider picture: analysing the system as a whole

Sharda et al., (2001) propose a framework to explain the technological and behavioral requirements in a laboratory environment. They suggest learning technologies should be categorized according to three dimensions: *temporal* (asynchronous/synchronous), spatial (colocated/ distributed) and level of presence (classroom/laboratory). They visualize this typology in the shape of a cube. This cube coincides with the taxonomy for experiments envisioned by COLDEX (figure 7). In the taxonomy the three axes are Location, Temporal and Nature of Experiment.

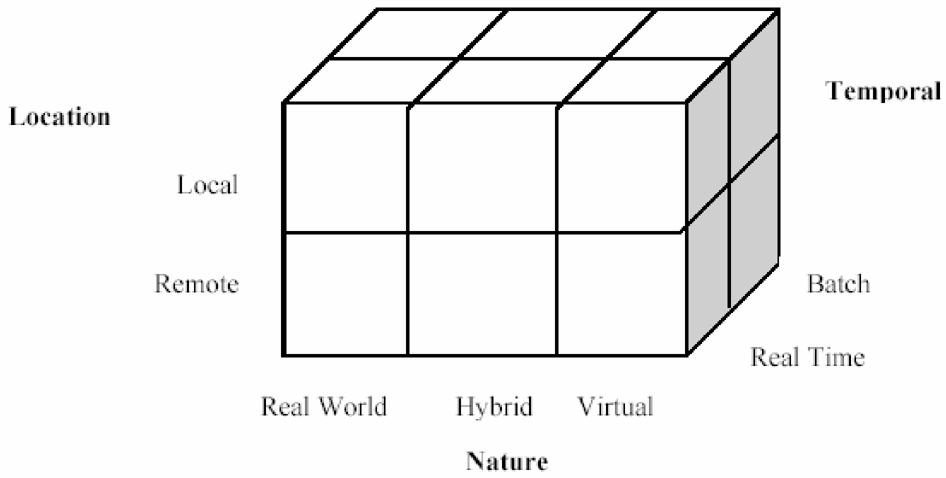

Figure 7. The taxonomy of experiments

An interesting way to look at the technological tools is to see them as boundary objects (also known as common information space). For example the tools are objects meant to be used by all members of OUS, but the users will use these boundary objects in very different ways. The concept of boundary objects was created by Star (1989) and may be explained as knowledge artifacts used in knowledge exchange. The boundary objects can be both carriers and containers. They carry meaning, shared interpretation or representation of an artifact, across communities' boundaries, which in the OUS also mean different locations, culture and time.

5.2.2 Evaluating tools and representations for learning

Human-Computer Interaction (HCI) is concerned with the design and evaluation of computer tools. The benefit of evaluating computer-based tools from the end user perspective is now well established. The goal is to design interactive computer systems that are effective, efficient, easy and pleasant to use. However, different concurrent conceptual perspectives exist that imply a distinct planning of the evaluation and methods to be utilized.

Historically, the discipline of HCI was strongly influenced by cognitive psychology. The idea was to bring cognitive theories to the development of human adapted computer systems. More specifically, the proponents considered that the theory should be able to: (a) provide sound support on the early stages of user interface design through the detailed description of the human cognitive characteristics and establishing clear implications on the way a certain product would be used and (b) provide conceptual and methodological frameworks to evaluate the user interfaces (Carrol, 1997, 2003). Yet, the growing need to study groups at work, develop computer based tools to support their activity and fundamental divergences regarding the conceptualization of cognition (for example, the mergence of situated action) led to a substantial enlargement of the theories and disciplines contributing to HCI. In particular, one should mention theories and conceptual frameworks that emphasize context and social organization aspects (Carrol, 1997, 2003; Olson & Olson, 2003) as, for example, Activity Theory (Nardi, 1996) and Distributed Cognition (Hutchins, 1995; Perry, 2003). Since Activity Theory has already been covered in previous sections we will briefly present the Distributed Cognition perspective and discuss its relevance for the COLDEX evaluation effort. Hollan, Hutchins and Kirsh (2002) consider that Distributed Cognition intends to go beyond solo cognitive activity and include the complex interactions between people, resources and the environment. It is an extension of the traditional cognitive science postulates and analysis that emphasizes the impact of social and environmental properties on cognitive processes. Hutchins, Hollan and Kirsh (2002) state two important principles of this perspective:

- The boundary of the unit of analysis - a cognitive process is delimited by the functional relationships between the elements of a certain phenomena and it is not constrained by spatial or even temporal boundaries.
- The cognitive mechanisms under consideration - the cognitive mechanisms include not only the manipulation of symbols inside the human agent but also the dynamic interplay between internal and external representations and symbol manipulation.

Salomon (1993b), focusing on the educational domain, thoroughly discusses the implications of adopting theories and conceptual frameworks that stress the importance of social, representational and environmental issues in cognitive phenomena. Salomon (1993a), however, stresses the need of not loosing sight of the individual cognitions since some cognitive phenomena in the educational contexts are inherently "solitary".

Finally, in this brief conceptual discussion, we need to focus on the representational level or, in other words, the representations that are used to learn. The relevance of this final comment is related to the fact that this level of analysis can be separated from the tools that convey them. It is a sub-specification of our digression through distributed cognition themes.

According to Scaife and Rogers (1996), there has been a focus on the internal mechanisms that act upon the external representations and not enough consideration has been paid to the dynamics between the external and the internal. Their external cognition framework seeks to fill this gap and focuses on the explanation of the dynamical relationship between internal and external representations. The paper addresses the question of determining which representations to use in learning environments. The authors' strategy to tackle the problem of conceptualising the relationship between the internal and the external involves the identification of the relevant cognitive properties of the external representations. They propose the following properties/notions:

- Computational offloading - relates to the phenomena of how informationally equivalent representations might require different degrees of cognitive effort.
- Re-representation - refers to the fact that problem solving can be eased by different ERs that share the same abstract structure.
- Graphical constraining - points out that some graphical components of a graphical representation are able to constrain the number and types of inferences made about the represented concept.
- Temporal and spatial constraining - relates to the possibility of making salient certain processes of a concept through the use of specific representational elements or strategies distributed over time and space.

Additionally, the authors propose the following design principles that can also frame specific research questions (Rogers & Scaife, 1998; Scaife & Rogers, 1996):

- Explicitness and visibility - different types of graphical ERs can make distinct aspects of the information encoded more relevant and/or are able to express more vividly underlying phenomena of complex processes.
- Cognitive tracing and interactivity - the issue here is the possible benefit of a user/learner being able to leave visible traces on the representation. The authors consider that this can imply a trade off between more or less interactive representations. More interactive representations although, for example, allowing testing and feedback, might make tracing more difficult.
- Ease of production - this design principle suggests the possible link between the ability to produce a certain type of representation and the ability to comprehend it more fully.
- Combinability - puts the emphasis on the possibility of combining different types of ERs. In this case, the designer must know not only the strengths and weaknesses of the different types of representations taken separately, but also the advantages and disadvantages of the combination of representations. The overall effect might not result from a simple additive process.
- Distributed graphical representations - in this case the issue lies on the possible differences that different contexts and social environments can impose, demand or facilitate on the use of ERs.

Why was this theoretical prelude necessary? Because it provides a coherent transition regarding conceptual and methodological issues from the wider macro perspective referred to in section 5.2.1 to the micro level that focuses on the tools and conveyed representations. In other words, this type of conceptual framework allows the structuring of possible "experiments"[1] at a micro level (concerning tools and representations) incorporating the impact of social and environmental properties and characteristics.

On the practical side, let us now look at the evaluation methods commonly used in HCI. Dix, Finlay, Abowd and Beale (2004) divide evaluation methods in two distinct groups: (1) the evaluation done through expert analysis and (2) the evaluation done through user participation. In the first group, the following four approaches are considered (Dix et al., 2004):

- *Cognitive walkthrough* - with this method a cognitive expert goes through the steps that an end user needs to take to carry out a specific activity looking for usability problems. The main focus has been to understand to what extend does the system supports exploratory learning. Users seem to prefer to explore the systems with hands on experiences than through complex instructions or manuals. To carry out a cognitive walkthrough one needs to: have a specification of the system; a description of the task to be pursued; a complete list of the actions needed to perform a task; an indication of the target user group.
- *Heuristic evaluation* - as the name indicates, this evaluation method uses heuristics to guide the design. In this context, a heuristic is a general design principle usually derived from basic research - it is an effort to put fundamental research into practice. This method is considered to be flexible and relatively cheap to implement. The idea is to have a group of evaluators checking a system independently for usability problems using a certain set

[1] Experiments are considered in a wide sense and not only highly controlled laboratory settings.

of heuristics. Research suggests that teams of 3 to 5 evaluators can spot 75% of the overall usability problems of a system.

- *Model based evaluation* - this approach is usually related to the use of cognitive models that have incorporated issues concerning design specifications and evaluation. For example, GOMS models are able to predict user performance given a certain structured and well defined task and interface.
- *Using previous studies* - in this case the idea is to review fundamental experimental research (for example, from experimental psychology) to inform/justify a particular design decision based on the results. It is, of course, extremely sensitive to a careful analysis of validation, especially ecological validity.

In the second group, one can find the following approaches:

- *Experiments* - controlled experiments can obviously provide strong evidence regarding design aspects. However, it is a time consuming and expensive approach that many times is not practical under a system development project. Having said that, it is unquestionable that some persistent design problems or phenomena can, and should, be studied under controlled conditions. The issue is to maintain a nice balance by taking into consideration the relevance of the effect to be found, the degree of generalization and ecological validity.
- *Observational techniques* that can be further sub-divided into:
- *Think aloud and cooperative evaluation* - on these evaluation methods the users are asked to verbalize its decisions and thoughts regarding and during the interaction with the system. However, care should be put concerning the effect that the verbalization process has on the task itself. The success of the approach is heavily dependent on the way the data is collected and analyzed. Furthermore, in order to counterbalance the effect of verbalization, the methodological approach can incorporate other types of observations in order to compare the differences on the interaction.
- *Automatic and non-automatic protocol analysis* - the emphasis here is on the recording of users' activity using a range of different techniques for posterior analyses. As examples: notes, audio recording, video recording, computer logging. The analysis can be more or less automatic depending on the degree of specificity a researcher can put on the range of events to look for and the type of analytical tools he is committed to (for example, tagging video recordings or computer logging and using statistical tools to look for users' patterns of interaction).
- *Post-task walkthroughs* - this method asks the users to reflect about their activity during the interaction with the system after the event. The user is presented with a record of his activity and is asked to comment it or answer questions regarding particular actions. It seems to be particularly useful in conjunction with the thinking aloud technique. For example, the user can be asked to perform two equivalent tasks, one using the think aloud method and the other post-task walkthrough. The analysis of the two tasks using the two methods can complement each other.
- *Query techniques* can be further sub-divided into:
- *Interviews*
- *Questionnaires*
- *Monitoring physiological responses* - recent research focused on the possibility of registering physiological measures during the interaction in order to "directly" assess aspects of computer use. As examples: eye tracking, heart tracking, activity of sweat

glands, and electrical activity on the muscle or brain. The fundamental issue, however, is to be able to find meaningful patterns and connect these to the intentional states of the user.

Summarizing, at this level of analysis, the choice of methods is wide and there are fundamental conceptual implications if one adheres to a single perspective. However, we believe that it is possible to use distinct methods as long as the overall conceptual framework that guides COLDEX is not lost of sight. The "triangulation" of different findings using different methods can be enriching.

6 Overview of data collection and analysis methods

While evaluating aspects within a COLDEX learning activity the aspect of focus is influencing the type of data collection. The following table (table 1) shows examples on focus, input data collection and analyzing methods. The evaluation methods may be both qualitative and quantitative. Using qualitative processes could include observation, videotaping, interviews and think-aloud narrative. Quantitative methods could be to use questionnaires, or to use process or network log data that the learning environment store automatically about the activities.

Table 1. Data collection and analysing methods.

Input Data	Analysis	Example	Description
Video	-Content-analysis (Semiotics), - Discourse analysis (Semiotics), - Video protocol analysis, - Video observational analysis - Interaction Analysis - Performance Analysis	- Peer to peer collaboration - Teacher – Student Interaction - Student – environment - Analysis of attitudinal behaviour using video data obtained from reciprocal settings	Video Data are crucial when focusing on the use of technological tools within a learning setting. The ability to capture the interaction of the learner with an object reveals possible downfalls and ways of improvement. The capturing of collaboration behavior between and among peers allows to accurately evaluating a learning environment or a learning activity.
Learning products (artefacts)	- Content Analysis (Semiotics), - Hermeneutic analysis - Structural Analysis	- Analysis of Cool Modes system dynamic student produced model with a focus on higher order thinking skills - Analysis of Student written reports to reveal scientific reasoning skills	Artefact such as written documents and software mediated artefacts should be used when the evaluation focuses on technological aspect or cognitive aspects. Student produced artefacts serve then as mental model representations that help to expose thought processes.

Process log	- Integrated analysis, - Feedback mechanism	- Logging data produced by the tools, e.g. Cool Modes	Barros and Verdejo (2000) use integrated analysis and feedback mechanism in their collaboration platform to provide qualitative description of group activities in relation to each member and to other groups.
Network log	- Social Network Analysis (SNA)	- Log data produced by the COLDEX network are used for analysis to find clusters and communities of practice - Measure relationships and flows between individuals, groups, and tools	SNA provides means to visually and a mathematically analyse human relationships. The humans are the node and the relationships are the links. These measures help determine the prominence of a node and the structure of the relations in the network.
Interviews (open-ended or semi-structured; structured)	- Discourse Analysis - Conversation Analysis - Inference and evidence	- To evaluate or assess a person or a group - To assess believes about facts, identify feelings and motives	Interviews may elicit reasons and explanation; past experiences; constructions of claims, activities, concerns; and project into the future (Cohen & Manion, 2000).
Questionnaire	- Content Analysis - Factor Analysis - Cluster Analysis	- To find students' or teachers' perception of something - To assess students' performance or behaviour - For students' self-assessment - To evaluate tools	Can be used to collect both qualitative and quantitative data. Use to find persons' opinions about an activity; some conditions; or relationships between specific events.

Observations – direct or participant (could be combined with note taking, parallel logging with video/audio recording, informal interviews)	-Analytic memo - Interaction Analysis -Ethnography - Time-Series Analysis -The analysing goes often hand in hand with fieldwork and interpretation most often takes its final form as a text.	- For case or field study - The evaluator is observing everyday practices -Participate in the ongoing events in a community -To get the teachers' and students' point of view -Try to understand how students and teachers work and get a picture of the context.	There could be different focus, for example event focus, person focus, place focus and object focus. Different research questions require different observational strategies. The goal is to build an understanding of the environment/context/ prerequisites and the people involved, teachers/students/etc, and how they interact with each other, tools and the environment. This will be a fundament for evaluating the DExT and the phenomena of learning when using a DExT.
Think-aloud	- Grounded Theory - Discourse Analysis	- Student task analysis - System evaluation	It is a way to capture the mental processes, the persons are asked to describe their thoughts and actions, to reflect on why and how choices are made.

In the following lines and to conclude this section, we describe in more details a number of analysis methods that have been presented in the table above.

There are different inquiry methods that can be used, such as Contextual Inquiry and Ethnography influenced field studies also called micro-ethnography (LeBaron & Streeck referred to by Stahl, 2003). The inquiry method that is used in field studies includes the following possibilities of data collection and analysis as appropriate for evaluating COLDEX activities:

- Informal, open-ended interview of users about their work, the way they use the COLDEX system, experiences, proposals, motivations, and opinions
- Observation: observing people using the COLDEX system in the way they do in everyday contexts. This often results in a descriptive understanding of the observed behavior. One reason is to see what users do to make inferences about the learning activity; a second is discovering what they cannot articulate in interview situations, or things of which they are not consciously aware.
- Notes and audio/video recordings are taken as memory aids, for discussions of events with the informant, to inform project members not involved in the field study, and also for later analysis and interpretation.
- Analysis and interpretation: In this work several project group members can be involved than those carrying out the field visits. In presenting the material, discussions may take place and members with "new eyes" can add questions and insights to the collected material.

Participatory Design is another very useful method where the user is viewed as one part in the design work. Common to these is an interest in the reality and the context where the studied phenomena occur, in our case where the end user is dealing with the COLDEX system. We need to follow the end user closely and take an interest in the whole context to discover not only what is obvious but more interestingly, the less obvious matters. We will find out how "people *actually* behave, not how they ought to behave" (Blomberg et al., 1993).

There are quantitative methods that may be a good alternative. One method could be to use questionnaires. Another method could be to use log data that the learning environment store automatically about the activities. Barros and Verdejo (2000) use integrated analysis and feedback mechanism in their collaboration platform to provide qualitative description of group activities in relation to each member and to other groups.

Lally & De Laat (2003) report the use of computer assisted data analysis software for Content Analysis where they divide the complex original messages in a networked collaborative learning environment, into meaningful units and assign a code to each unit, to be able to extract, generalize and abstract. The codes are e.g. knowledge construction processes and social processes with metacognitive codes that the users used to mark questioning, expressive, explaining, and share ideas. In the next section we will provide you with an overview of methods for data collection and analysis that can be relevant for COLDEX scenarios.

Video recording is sometimes used for narrative or conversation analyses. This is a quite time consuming method. Bødker (1996) describes how her research group analyzed a four hour videotaped interaction while combining ethnography and interaction analysis. The ethnography was important for the understanding of the contextualization. The video contained an event log with a description and chronological index of the events. By using the concepts of breakdown and focus shifts they gathered instances of particular interest. They then selected four interesting sections of a few minutes' length each. Secondly, they mapped the action by (a) listing the objects that the user focused on and (b) listing the narrative of the situation including notes of what the physical actions of the user. The breakdown and focus shifts were used to provide pointers for understanding how the artifacts mediate activity.

7 Action Plan for the COLDEX evaluation

The previous chapters in this deliverable have laid a theoretical foundation for the evaluation activities that have to be done within the COLDEX project. Each partner has the possibility to choose and evaluate certain aspects relevant to the learning activities to be conducted at each site. There are though, some common parts that every partner responsible for a learning scenario have to address. This chapter gives more tangible directions on how that common evaluation shall be conducted. The directions are intended for the COLDEX partners in Sweden, Germany and Spain. The targets for the evaluation are the different learning scenarios that each partner has constructed, i.e. the BioDiversity, Astronomy, the maze and the Chemistry scenario. The interrelation between the different COLDEX scenarios will be evaluated by looking at the Open User Scheme, the DexTs, and the LOR.

7.1 Learning requirements

The COLDEX project supports the notions of challenge-based, collaborative learning and distributed experimentation. Some of the overall evaluation questions derive from the COLDEX learning requirements (Deliverable 2.2.1). They should guide the evaluation activities to be carried out by each partner, especially in the field studies. The partners are asked to assess how their learning activity correlates with the requirements. These specific aspects are further explored by the specific questions presented below. At the end of this section the reader can find a checklist that can be used after the specific questions have been answered.

Authentic activities (also situating the context): Authenticity within the field of science learning can be described in various ways. As described in D 2.2.1, COLDEX aims to develop authentic environments by providing a highly contextual setting, which allows learners to learn science in the environment in which it actually occurs and to learn science by using authentic scientific methods. Specific questions to be asked:

- *To what extent does the activity /system support authentic scientific inquiry activities?*
- *Are the activities situated in a rich context?*

Construction: Students are encouraged to not only engage in hands-on but also in minds-on activities by having the opportunity to build artifacts collaboratively and to share them. Specific questions to be asked:

- *To what extent does the system allow the building of artifacts like e.g. scientific journals?*
- *To what extent do the students engage in this activity and when are they doing it ?*
- *Can artifacts be uploaded, viewed and shared in the LOR?*
- *Are the artifacts viewed and shared?*

Collaboration: Collaboration when viewed as an educational method has various benefits within science learning and experimentation. Specific questions to be asked:

- *In what ways and when does collaboration take place?*
- *How well is collaboration supported system wise and pedagogically?*

Reflection: Reflection is especially important when investigating complex phenomena.
Specific questions to be asked:

- *Do the activity and the environment allow that the students apply strategic procedures to solve scientific problems?*
- *To what extent does the activity foster reflection?*
- *Do what extent does the system add value to student learning*

Multi-modal interaction: Every learner has his or her own way to learn and needs different kinds of support. To have opportunities to use “multi-modal interaction” is one way to solve the problem. Specific questions to be asked:

- *To what extent does the system support multi-modal interaction?*
- *To what extent does the system support interoperability ?*

Table 1 is a check box that can be filled in to sum up how the learning activity meets the requirements specified above.

Table 1. Requirements for a Learning Activity.

Requirements for Learning activity	**Included Yes/No**	**How/ what?**
Authentic activities		
Construction of artifacts		
Collaboration		
Reflection		
Situating the context		
Multi-modal interaction		

7.2 The learning environment

A simple way to assess the learning environment would be to use the checklist in table 2 *"Learning Environment Requirements"* based on D.2.2.2. The purpose of this table is to see how the evaluated scenario meets the learning requirements.

Table 2. Learning Environment Requirements.

Requirements for Environment	**Included Yes/No**	**How/ what?**
Support for the collaborative construction of knowledge objects		
Tools that support negotiation		
Both public and private feedback		
Mechanism to share and exchange information, objects, views, etc.		
Facilitate a meaningful division of labor		
Support joint online thinking, commentary		
Foster a sense of collaborative learning community		
Support mediation among all the participants		
Integration of online and offline, individual and collaborative, in-class and distributed activities		
Flexibility to adapt the environment to the local conditions (students' background and capabilities and/or teacher's preferred teaching style)		

7.3 Framing the activity

In order to describe the context in which the learning activity is situated in, activity theory can be used as a guide to structure the description (see section 3). That activity shapes an activity system with nodes or entities, and relationships between the entities. The following table (table 3) could be useful when outlining the activity system or maybe even a network of activity systems. The first column names the entity or node, the second suggests a question to help you identify that node. The last column named “Answer” has to be filled with the appropriate answer connected to the question in the previous column.

Table 3. The entities of an expanded activity system.

Entity	Question	Answer
1. The subject	*Which individual’s or group’s viewpoint will be used?*	
2. The object of activity	*What is the goal or motive that will be acted upon?*	
3. Tools or mediating artifacts	*What mediating instruments are used?*	
4. The community of practice	*Who else share the same object?*	
5. Division of labor or roles	*Who does what?*	
6. Rules	*What rules of conduct, regulations or constraints exist?*	
7. Outcome	*What will be the result or product? What will the object be transformed into?*	

An example of how activity theory can be applied to give a rich description of the background and context of a learning scenario can be found in D 8.1.1. (section BioDiversity scenario) This description also should contain references to other deliverables, information about audience, content domain, goals and purpose.

7.4 Evaluating learning aspects

The aspects of an evaluation that is described in this section, should be used to guide the individual scenario partners to meet the overall aspects of an evaluation. They will be described in more detail below together with instructions on what instrument to use. The field study should include at least a one-to-one trial with one or many users. The criteria for this trial should be selected by each partner according to the learning goal and the instructional goal of the scenario.

The table below (table 4) gives an overview of how relevant variables related to COLDEX can be assessed using appropriate instruments. Below the table recommendations will be given which instrument is most feasible to use with in the COLDEX evaluation.

Table 4. Methods (adapted from Cohen and Manion, 1989).

Variables→ Instruments ↓	Knowledge and comprehension	Attitudes and motivation	Transversal skills, specifically: Collaboration	Science Skills	Usability	Language issues
Questionnaire	*	*	*	*		
Scale ratings		*	*	*		

Observation			*			*
Heuristic evaluation					*	
Standardized tests /performance assessment	*					
Interview, Focus Groups			*			*
Logfiles			*		*	*
Learning diary		*				

- *Knowledge and comprehension*

In order to assess knowledge and understanding acquired within the COLDEX scenarios, assessment items need to be developed according to the topic taught within the individual scenarios. For several topics TIMMS assessment (Mullis et al., 2003) or SOLO-taxonomy (Biggs & Collis, 1982) can be used as a guide to adapt an assessment instrument for assessing knowledge and comprehension.

- *Attitudes and Motivation*

In order to assess motivation related behavior assessment items need to be developed that can be included in a questionnaire. These items can be created in alignment with intrinsic motivation inventory by Salmi (2003). The intrinsic motivation inventory takes into account situational behavior and is developed for learning in informal learning settings.

Additionally to motivational issues attitudinal behavior towards technology in general can be assessed by using items adapted according to the snapshot survey for teachers developed by Norris and Soloway (1998).

- *Transversal skills, specifically: Collaboration*

Collaboration is one of the learner requirements of COLDEX. Derived from the evaluation questions of the action plan, the learning requirements rating scale can be used within an expert observation to assess the level of collaboration that can take place within a certain learning environment.

- *Science Skills*

Science skills as specifically described in the framework section of the evaluation plan, can be evaluated either by observing or interviewing the learner, but also by the level of support of science skills within the environment. Malhotra and Chinn (2002) provide a useful rating scale to assess authentic science skills within scientific inquiry activities. It can be used if the overall structure of scientific inquiry is encouraged to be used within a science activity. It can be either used the rating scale template developed by Chinn and Malhotra to assess cognitive processes supported within authentic inquiry tasks or the check box survey developed to by Chinn and Malhotra to assess the authenticity within inquiry tasks.

- *Usability*

One big part within usability evaluation needs to be the overall effectiveness of computer-based learning systems in particular the LOR. To evaluate the system a heuristic evaluation from the experts' perspective is recommended as well as a one trial user evaluation and focus groups. Templates developed by Reeves (2003) can be used, but need to be adapted to the LOR functionality. Suggestions for adaptation and a draft template for carrying out a heuristic evaluation should be included in evaluation plan II.

A second user trial evaluation should be carried out to assess how well the objectives of the individual project were met. The objectives can be assessed in two ways depending on what kind

of information is available. A post test can be developed in alignment with the learning objectives that allows to assess what the students learned.
The usability according to teacher acceptance can be assessed by conducting interviews within teacher focus. Questions should include: Is it possible for the teacher to implement scenario activities independently in the classroom? Does the scenario contribute to the curriculum or does it enhance the curriculum?

7.5 Evaluation instruments

This is a list of instruments, which should be used for COLDEX evaluation. They will be adapted and enhanced with additional context specific assessment items.

TIMMS Test:

http://timss.bc.edu/TIMSS1/Items.html

SOLO-taxonomy:

http://www.shef.ac.uk/nlc2002/proceedings/symp/10.htm

Teachers' Attitude toward technology survey developed by Norris and Solloway (1998). (15 min)

http://www.tcet.unt.edu/research/survey/snapshot.htm

Heuristic evaluation, Focus group, user testing etc.:

http://it.coe.uga.edu/~treeves/edit8350/tools.html

Authentic scientific inquiry rating scales in:

Chinn, C. A., & Malhotra, B. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. Science Education, 86(2), 175-218.

http://www3.interscience.wiley.com/cgi-bin/abstract/90513046/ABSTRACT

Motivation developed by Hannu Samli (2003) in:

Salmi, H. 2003. Science centres as learning laboratories: experiences of Heureka, the Finnish science centre. International Journal of Technology Management. Vol. 25, No. 5,pp.460-476.

7.6 Logistics /Timeline

A more detailed plan of how and when to conduct the evaluation at each COLDEX site will be specified within the evaluation plan II (WP8). Table 5 and 6 show how such schedule and timeline may look like.

Table 5. Schedule of evaluation (adapted Reeves, 2003)

Question	Actual
Who will implement the evaluation plan and collect the data	
Which data will be collected	
How will you analyze the effectiveness of your scenario?	

Table 6. Timeline

Activity	Time
Evaluation planning and definition of a specific task	October - November
Conducting evaluation	November - February
Analyzing data and writing the final evaluation report	March - May

References

Anderson, N. J. (2002). *The Role of Metacognition in Second Language Teaching and Learning*. ERIC Digest. Education Resources Information Center. Retrieved from http://ericadr.piccard.csc.com/extra/ericdigests/ed463659.html

Barab, S. A., Barnett, M., Yamagata-Lynch, L., Squire, K. & Keating, T. (2002). Using Activity Theory to Understand the Systemic Tensions Characterizing a Technology-Rich Introductory Astronomy Course. In *Mind, culture, and activity,* 9(2), pp 76-107.

Barros, M., and Verdejo, M. (2000). Analysing student interaction processes in order to improve collaboration. The DEGREE approach. *International Journal of Artificial Intelligence in Education, 11*, 221-241.

Bereiter, C. (2002). *Education and Mind in the Knowledge Age.* NJ: Lawrence Erlbaum.

Biggs, J. and Collis, K. (1982) .*Evaluating the Quality of Learning: the SOLO taxonomy.* New York: Academic Press.

Blomberg, J., Giacomi, J., Mosher, A., Swenton-Wall, P. (1993) Ethnographic Field Methods and Their Relation to Design. InD. Schuler and A. Namioka. *Participatory Design Principals and Practices.* NJ: Lawrence Erlbaum.

Bloom, B.S. (Ed.). (1956). *Taxonomy of educational objectives: The classification of educational goals: Handbook I, cognitive domain*. New York ; Toronto: Longmans, Green.

Bødker, S. (1996). Applying Activity Theory to Video Analysis. In B. A. Nardi (ed.), *Context and consciousness. Activity Theory and Human-Computer Interaction.* 2nd printing, 1997. Cambridge, MA: The MIT Press.

Carrol, J. M. (1997). Human-computer interaction: psychology as a science of design. *International Journal of Human-Computer Studies, 46*, 501-522.

Carrol, J. M. (2003). Introduction: Toward a Multidisciplinary Science of Human-Computer Interaction. In J. M. Carrol (Ed.), *HCI Models, Theories and Frameworks* (pp. 1-10). San Francisco: Morgan Kaufmann Publishers.

Chan, C. K. K. & Van Aalst, J. (2003). Assessing and Scaffolding Knowledge Building: Pedagogigal Knowledge Building Principles and Electronic Portfolios. In B. Wasson, S. Ludvigsen, & U. Hoppe (eds.). *Designing for Change in Networked Learning Environments. Proceedings of the International Conference on Computer Support for Collaborative Learning 2003.* Dordrecht, NL: Kluwer Academic Publishers.

Chinn, C. A., & Hmelo-Silver, C. E. (2002). Authentic inquiry: Introduction to the special section. Science Education, 86(2), 171-174.
-->http://www3.interscience.wiley.com/cgi-bin/abstract/90513044/ABSTRACT

Chinn, C. A. and Malhotra, B.A. (2002) Epistemologically authentic inquiry in schools: a theoretical framework for evaluating inquiry tasks. *Science Education*, 86(2), 175-218.
--> http://www3.interscience.wiley.com/cgi-bin/abstract/90513046/ABSTRACT

Christiansen, E. (1996). Tamed by a Rose: Computers as Tools in Human Activity. In B. A. Nardi (ed.), *Context and consciousness. Activity Theory and Human-Computer Interaction.* 2nd printing, 1997. Cambridge, MA: The MIT Press.

Cohen, L & Manion, L (1989) Research Methods in Education, London, Routledge

de Jong, F. P. C. M., Veldhuis-Diermanse, E., & Lutgens, G. (2002) Computer supported collaborative learning in university and vocational education. In T. D. Koschmann, R. Hall & N. Miyake (Eds.), *CSCL 2: Carrying Forward the Conversation.* Mahwah, NJ: Erlbaum.

Dix, A., Finlay, J., Abowd, G., & Beale, R. (2004). *Human-Computer Interaction*. Harlow: Pearson Education Limited.

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

Germann, P.J., Haskins, S., & Auls, S. (1996). Analysis of nine high school biology laboratory manuals: Promoting scientific inquiry. *Journal of Research in Science Teaching, 33 (*5), 475-499.

Hofstede, G. (2001). Culture's consequences : comparing values, behaviors, institutions, and organizations across nations, 2nd ed. Thousand Oaks: Sage.

Hollan, J., Hutchins, E., & Kirsh, D. (2002). Distributed Cognition: Toward a New Foundation for Human-Computer Interaction Research. In J. M. Carrol (Ed.), *Human-Computer Interaction in the New Millenium* (pp. 75-94). New York: ACM Press.

Hutchins, E. (1995). *Cognition in the wild*. Bradford: MIT Press.

Jerkedal, Åke (1999), *Utvärdering – steg för steg.* Stockholm: Norstedts Juridik AB

Kaptelinin, V. (1996). Activity Theory: Implications. In B. A. Nardi (ed.), *Context and consciousness. Activity Theory and Human-Computer Interaction.* 2nd printing, 1997. Cambridge, MA: The MIT Press.

Koschmann, T. (1996), Paradigm Shifts and Instructional Technology: An Introduction. In T. Koschmann (ed.) *CSCL: Theory and Practice of an emerging paradigm.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Kuuti, K. (1996), A framework for HCI Research. In B. A. Nardi (ed.), *Context and consciousness. Activity Theory and Human-Computer Interaction.* 2nd printing, 1997. Cambridge, MA: The MIT Press.

Lally, V & de Laat, M. (2003) A Quartet in E. In B. Wasson, S. Ludvigsen, & U. Hoppe (eds.). *Designing for Change in Networked Learning Environments. Proceedings of the International*

Conference on Computer Support for Collaborative Learning 2003. Dordrecht, NL: Kluwer Academic Publishers.

Law, N., & Wong, E. (2003). Developmental Trajectory in Knowledge Building: An Investigation. In B. Wasson, S. Ludvigsen, & U. Hoppe (eds.). *Designing for Change in Networked Learning Environments. Proceedings of the International Conference on Computer Support for Collaborative Learning 2003.* Dordrecht, NL: Kluwer Academic Publishers.

Mullis, I.V.S., M.O. Martin, T.A. Smith, R.A. Garden, K.D. Gregory, E.J. Gonzalez, S.J. Chrostowski, and K.M. O'Connor (2001). TIMSS Assessment Frameworks and Specifications 2003. Boston, MA: Boston College, The International Study Center.

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

Nash, John; Plugge, Leo; & Eurelings, Anneke (2001), Defining and Evaluating CSCL Projects, managing towards evaluation. Paper presented at *Euro-CSCL 2001,* Maastricht, Netherlands.

Nurmela, K., Palonen, T., Lehtinen, E. and Hakkarainen, K. (2003). Developing Tools for Analyzing CSCL Process. In B. Wasson, S. Ludvigsen, & U. Hoppe (eds.). *Designing for Change in Networked Learning Environments. Proceedings of the International Conference on Computer Support for Collaborative Learning 2003.* Dordrecht, NL: Kluwer Academic Publishers.

Olson, G., & Olson, J. (2003). Human-Computer Interaction: Psychological Aspects of the Human Use of Computing. *Annual Review of Psychology, 54*, 491-516.

Ossimitz, G. (2000). *Entwicklung systemischen Denkens*. München: Profil Verlag.

Perry, M. (2003). Distributed Cognition. In J. M. Carrol (Ed.), *HCI Models, Theories and Frameworks* (pp. 193-224). San Francisco: Morgan Kaufmann Publishers.

Preece, J.; Rogers, Y. & Sharp, H. (2002). *Interaction Design, beyond human-computer interaction.* NY: John Wiley & Sons, Inc.

Reeves, T. C., & Hedberg, J. G. (2001). Effectiveness evaluation. Impact evaluation. Interactive learning system evaluation [Electronic version]. Retrieved October 29, 2004, from the University of Georgia, Department of Instructional Technology web site: http://it.coe.uga.edu/~treeves/evalbook/eilsbook.htm

Rogoff, B. (1995), Observing sociocultural activity on three planes: Participatory appropriation, guided participation, and apprenticeship. In J.V. Wertsch, P. del Rio, & A. Alvarez (eds.). *Sociocultural studies of mind*, Cambridge, UK: Cambridge University Press, pp. 139-164.

Rogers, Y., & Scaife, M. (1998). How can interactive multimedia facilitate learning? In J. Lee (Ed.), *Intelligence and multimodality in multimedia interfaces: Research applications.* Menlo Park, CA: AAAI Press.

Roschelle, J. (1996), Convergent Conceptual Change. In T. Koschmann (ed.) *CSCL: Theory and Practice of an emerging paradigm.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Salmi, H. 2003. Science centres as learning laboratories: experiences of Heureka, the Finnish science centre. International Journal of Technology Management. Vol. 25, No. 5,pp.460-476.

Salomon, G. (1993a). No distribution without individuals' cognition: a dynamical interactional view. In G. Salomon (Ed.), Distributed Cognitions: Psychological and educational considerations (pp. 111-138). Cambridge: Cambridge University Press.

Salomon, G. (Ed.). (1993b). Distributed Cognitions: Psychological and educational considerations. Cambridge: Cambridge University Press.

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? International Journal of Human-Computer Studies, 45(2), 185-213.

Scardamalia, M. (2002) *Collective Cognitive Responsiblitiy for the Advancement of Knowledge.* IKIT, OISE, University of Toronto.

Schoenfeld, A. H. (1987). What's all the fuss about metacognition? In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 189-215). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schuler, D. & Namioka, A.(1993). *Participatory Design Principals and Practice.* Hillsdale, NJ: Lawrence Earlbaum Associates, publishers.

Sharda, R., Romano, Jr., N.C., and Lucca, J. (2001). *A Conceptual Framework for Computer Supported Collaborative Learning Requiring Immediate Presence (CSCLIP).* Human Resource Development in a Networked World, Vol. 33, pp. 111-126.

Sins, P., Elwin S., van Joolingen (2003). Characterizing secondary students' dynamic modeling processes. Paper presented at the *European Association of Research in Learning and Instruction* Conference.

Smith, P. L., & Ragan, T. J. (1999). Instructional Design (2nd ed.). New York: John Wiley & Sons, Inc.

Soloway, E., Norris, C., Knezek, G., Becker, H., Riel, M. & Means, B. (1999) The Relationship of Teachers and Technology: Survey Findings and Reflections. Panel presented at Society of Information Technology & Teacher Education (SITE)'s 10th International Conference, San Antonio, TX, March 3, 1999.

Stahl, G. (2003) Meaning and interpretation in collaboration. In B. Wasson, S. Ludvigsen, & U. Hoppe (eds.). *Designing for Change in Networked Learning Environments. Proceedings of the International Conference on Computer Support for Collaborative Learning 2003.* Dordrecht, NL: Kluwer Academic Publishers.

Star, S.L. (1989). The structure of ill-structured solutions: boundary objects and heterogeneous distributed problem solving. *Distributed Artificial Intelligence*, 2, 37-54

Stratford, S. J., Krajcik, J., Soloway, E. (1998). *Secondary students' dynamic modeling processes: Analyzing, reasoning about, synthesizing, and testing models of stream ecosystems*. In *Journal of Science Education and Technology*, 7(3), 215-234.

Wertsch, J. V., del Rio, P. & Alvarez, A. (1995), Sociocultural studies: history, action and mediation. In J.V. Wertsch, P. del Rio, & A. Alvarez (eds.), *Sociocultural studies of mind*, Cambridge, UK: Cambridge University Press.