Journal of Universal Computer Science, vol. 4, no. 4 (1998), 349-381
submitted: 10/12/97, accepted: 28/12/97, appeared: 28/4/98 © Springer Pub. Co.

Bridging Qualitative and Quantitative Approaches in Evaluating the Educational Effectiveness of a Shared Design Memory

Daniela Giordano

Department of Education
Graduate Program in Educational Technology
Concordia University
1455 de Maisonneuve Blvd. W., Montreal, Quebec, H3G1M8, Canada
E-Mail: giord@alcor.concordia.ca

Abstract: A shared design memory emerging from the contributions of novice designers affords, theoretically, unique opportunities to support individual and organizational learning. Evaluation must take into account the „distributed" nature of the system that becomes realized. The proposed evaluation model is based on a cross-analysis of: the contents of the shared design memory, the quality of the design artifact produced be the designers teams, the characteristics of the student population, and their perceptions of the adequacy and usefulness of the representational formats adopted for the shared memory. Effects being sought are generational changes that indicate that design weaknesses typical of novices are being offset, and that good design practices are diffused and gradually incorporated as new quality standards. Preliminary results of the evaluation of a shared memory for Information Systems design show that shared memory underpinned an emergent quality in the new designs, characterized by increased structuredness, communicability, and attention to the dynamics of interactions in the system being designed. The shared memory was deemed useful and usable by the learners. Findings also clarify the relative merits of some representational formats (links among design cases and reviews attached to design cases) used for conveying design knowledge.

Keywords: shared design memory, organizational learning, information systems design, evaluation
Categories: K.3.1, K.3.2, H.5.1, H.5.3, H.2.1

1 Introduction

One emerging class of educational software comprises environments that are meant to facilitate learning by encouraging productive forms of collaboration among the learners and by making possible the creation of accessible, shared resources instrumental to the activities of a specific community of learners. Broadly, such environments may be seen as applications of Computer-Supported-Collaborative-Learning (CSCL) (e.g., [Silverman, 95]; [Koschmann, 96]).

The intermediate or final product of collaboration, e.g., the solution to a problem, or the argumentation towards building a shared understanding of a phenomenon, or a design artifact, can be recorded for later use by other groups or members of the community. In this case, the supporting software also aims at creating electronic „organizational memories", where the „organization" can be the classroom [Bruegge & Coyne, 94] or the virtual community of learners who have been enrolled in the same course in different years [Giordano, 97a]. Peculiarly, such learning environments, typically based on distributed hypertexts architectures augmented by communication and repository technologies, are „content-free“. They rather provide basic mechanisms affecting the structure of communication and the way resources for learning can be organized and accessed, thus encouraging (or discouraging) specific kinds of interactions and activities. For example, the learners' notes and discussion threads can be labeled according to types such as „the problem“, „my theory“, „evidence“, „I need to understand“ and so on [Scardamalia & Bereiter, 93]. On the other hand, such environments are „mode of discourse-dependent“ because the nature of the learning goals, the domain pertaining the competencies to develop (e.g., scientific, technical, design, literary discourse and reasoning) and the constraints of the instructional situation may decree the appropriateness and efficacy of one set of communication and content organization mechanisms over another one.

Although the design of collaborative systems is increasingly grounded on principles of learning mostly stemming from the cognitive tradition to motivate the design features of the systems [Koschmann et al. 96], and deployment is informed by the accumulated wisdom on how social actors interact, these theoretical foci are not sufficient to ensure that learning is taking place. Evaluation of these environments is challenging both on methodological and theoretical counts [Brown, 92], because their deployment amounts to a complex intervention that largely affects the space of educational affordances and is often associated to emphasis on worthwhile but elusive learning goals such as fostering reflection, communication, articulation of ideas, collaboration, motivation, or enhancing metacognition.

Separability of the effects and their tracing to specific features of the technological environment and of its underlying pedagogical model is further complicated by the inherently social and distributed nature of the system that becomes realized, as it is accounted for in the paradigms of situated and distributed cognition [Brown et al., 89]; [Lave, 88]; [Salomon , 93]; [Pea, 93].

Most of the evaluation studies have focused on rich descriptions of how the system is being used, with much emphasis on its educational affordances, mostly following the case study method; the few quantitative studies that have been done have focused on the relationships between types of usage pattern and individual conceptual progress [e.g., Oshima et al., 96]. However, the relationship between the individual dimension of learning and the social, or „organizational“ one, seems always to be left in the background as an unresolved issue, especially when it comes down to defining and evaluating the efficacy of the learning environment.

The article focuses on the approach that is being used to evaluate the instructional leverage afforded by a shared electronic design memory which is emerging from cycles of contributions of communities of learners engaged in Information Systems analysis and design [Faro & Giordano, 97a]; [Giordano, 97b]. The system allows the students to retrieve precedent design cases developed by other students, possibly for reuse and adaptation to their current design problem, and to examine (if any) the peer reviews, or any other comment, that have been made to such design cases.

The proposed evaluation method has at its core the relationship between individual learning and the learning of the community that hosts and sustains the process. The latter is referred to as „organizational learning“ [Levitt & March, 88]; [Huber, 91]; [Brown & Duguid, 91], and accounts for the fact that although in such learners' communities there is a complete turn-over of the participants (typically with a yearly frequency), they still manage to transmit myths, procedures and expectations, as well as content knowledge and misconceptions that both enable and limit the performance of the organization. The leading evaluation questions are to what extent designers are facilitated in carrying out tasks notoriously difficult for novices and in what respects the quality of their analysis and of their design artifacts is affected by using and contributing to the shared design memory. The units of analysis are both individuals or small design teams and the community of learners as a whole.

The article is organized as follows. [Section 2] briefly addresses the rationale behind the educational use of a shared design memory. [Section 3] illustrates how the general architectural principles of a shared design memory have been actualized in a system to learn Information Systems Analysis and Design. [Section 4] presents the evaluation model, which is based on a cross-analysis of data originating from the contents of the shared design memory, the actual design artifacts produced by the teams, and from questionnaires to gather individual attitudes. [Section 5] presents and discusses the results concerned with the students' attitudes and the contents of the shared memory [Section 6], after illustrating the approach adopted to evaluate design quality completes the preliminary assessment by characterizing the emergent properties of designs produced under shared memory. Some concluding remarks are in [Section 7].

2 Shared Design Memory as a Learning Tool

Key to the problem of supporting organizational learning is finding ways to make the knowledge embedded in artifacts, documents, and pointers to human experts more resilient and more widely available to the potentially interested members in an organization, i.e., enhancing the organization's „memory" [Walsh & Ungson, 91]; [Giordano, 97a]. The informal social networks that are established among the members of the organization are an essential component of the organizational memory, by providing (or blocking) the necessary pointers to relevant expertise sources, and by fueling, mostly through a process of narrative construction, the creation of a shared culture and knowledge base anchored to practice [Brown & Duguid, 91]. In a learning context, the processes above partly concretize in the informal talks with peers who have undergone the same experience, in the exchange of resources (notes, references, etc.) and originate the expectations and understandings that are brought by the students in approaching a new course. The relevance of such informal component of the organizational memory also underscores the limits of what can be represented and „recorded" to make the organization less vulnerable to turn-over of key members, and in general, perform better.

Electronic support of the organizational memory can take many forms, ranging from maintaining an updated system of answers typically asked of the experts [Ackermann & Malone, 90] to case-based reasoning systems [Kolodner, 93], which allow the retrieval of the cases (similar to the problem at hand) with the relevant solution. Experimentation on recording the rationale for decisions, especially those occurring in design meetings, has generated research into languages and graphical notations for expressing the argumentation underlying the design decisions, by structuring the discussion that surrounds any particular design decision, according to categories such as „support“, „object“, „goal“ [e.g., Conklin & Begeman, 88]. On the one side, the choice of one representational language amounts to a commitment to a certain mode of discourse or to the assumption that it fairly represents the one that is used in practice. On the other side, later intelligibility of the rich resulting picture is difficult even for those who participated, for the lack of context or contextual clues [Buckingam Shum, 97]. This makes even more difficult the transfer or „reuse“ of that design experience across different projects or teams. The two points above highlight two requirements central to our discussion, i.e., that both what is recorded and the underlying representational format be as much targeted as possible in supporting the knowledge transfer to new generations of users (as the supervening students in a design course) to support continuity and enhance the process of acquiring and transforming design practices.

Part of the rationale for incorporating the process of using and building a shared design memory in the context of learning a design practice is that this process supports many modes of learning, such as learning by example, learning by collaboration, learning by critiquing, learning induced by the effort of clearly articulating ideas or formulating a query to search for a needed resource. The possibility of inspecting a variety of design artifacts is useful because such artifacts embed technical knowledge about how certain design cases are solved, the variety of cases can suggest more refined dimensions for analysis and help develop cognitive flexibility [Spiro et al., 91].

Most of the above modes of learning could be activated simply by using any resource that can provide enough richness in the examples, such as a ready-made, case-based system [Barber et al., 92], or a cognitive flexibility hypertext [Spiro et al., 91]. However, using and contributing to a shared organizational memory that is owned by the learning community in which it is used, because it gradually emerges from what has been done in the past and mirrors what is of current interest, provides some unique opportunities.

First, it is an experience that is „authentic“, al least in two respects: the learners are introduced to a process of knowledge construction in which they have to take personal responsibility, rather than taking knowledge as a given, and they exercise the communication skills that facilitate transferring and reusing experience, in the effort to making their contribution as intelligible as possible (even to those that did not share the context of its creation) [Giordano, 1997a]. This latter aspect is especially important in domains in which competence requires being able to participate in the collaborative efforts of an interdisciplinary team.

Second, especially in the domain of design (in which there is a marked creative and personal component both in the way of framing problems and of finding solutions) the fact that the design cases are shared provides a reference model of the quality that can be achieved. This can be a motivational drive for those learners that are truly novices to the field, and also for those who are more experienced and confident. Thus one element of educational leverage is provided by the means of creating individual „identities“ as a designer in the space of the shared design memory, which implies being able to relate and differentiate one’s own artifact from those that already exist [Giordano, 97a].

Because the contributions to the shared memory originate from novices, variation in their quality must be expected. This is an opportunity for the students to exercise critical skills and avoid „blind reuse“ of the design cases. It is also an opportunity to highlight design weaknesses in context and make them meaningful to a wider audience. But it is also entails the risk of promoting uniformity or the diffusion of incorrect or unproductive practices if the users do not generate enough variety and insights to make the system self-correcting. Thus, to fully gather the potential educational benefits, the adopted representation and structuring mechanisms must be carefully tailored to the design domain, ensuring consistency with the sought after modes of discourse, and must embed some features that counteract the risk of uniformity.

These issues are central to the context problem, i.e., preserving and enhancing the meaningfulness of the information, and to deciding to what extent product and process must be represented in the memory, for example by summarizing the process that led to the solution of a problem together with the solution. How the above issues have been taken into account in the architecture of a shared design memory in the domain of Information Systems analysis and design, is illustrated in the next section.

3 StoryNet: a Shared Design Memory for Information Systems Analysis and Design

StoryNet [Faro & Giordano, 97a]; [Giordano, 97b] is a shared memory for the introductory course to Information Systems (IS) Analysis and Design at the undergraduate level in an information engineering program. Broadly, Information Systems analysis and design involve gaining knowledge about the business organization, building a conceptual model and „transforming“ it into a formal representation detailing the requirements in a form suitable for design and implementation considerations. This transformation is very demanding because structured analysis and design methodologies, although providing procedural guidelines and diagramming notations, support only partially analysis, requirements acquisition and comprehension [Sutcliffe & Maiden, 92]. Capturing the requirements specifications in a conceptual model of the system requires solid analytical skills but is also affected by familiarity with the application domain, which underpins a sense of relevance in identifying the data and key processes in the organization. Mastering specific modeling techniques such as data modeling (e.g., the entity-relationship formalism), function modeling (e.g., data flows), process and event modeling (e.g., Petri networks) is a relevant component of IS analysis and design competence, but at its core are processes of continuous analysis and synthesis involved in the integration of different views, in tracing specifications to requirements, validating the semantic and dynamic aspect of the model under construction, and performing the ill-structured transition from the conceptual model to the architecture of the system to be implemented (e.g., in our case, a relational database).

StoryNet aims at supporting individual learning by enhancing the ability to carry out a deep, user-centered analysis of the business organization, soliciting uniqueness in the proposed design solution, and consolidating the ability to critique and verify the design, and organizational learning by supporting the circulation and acquisition of design ideas and practices. Its architecture stems from the idea of getting instructional leverage by jointly operating on an explicit model of the biases and difficulties of novice analysts and designers, and on the methodological approach to design.

The extant literature on the cognition of novice data modelers [Batra & Antony, 94] and systems analysts [Maiden & Sutcliffe, 92] highlights some recurrent difficulties and biases in carrying out the analysis and design process. The difficulties involve scoping the problem and recognizing its boundaries; performing problem decomposition; reasoning on model completeness; and generating and testing hypothesis about the model by robust problem-solving strategies (such as heuristic-based reasoning or use of scenarios). The biases are: the tendency to concentrate immediately on implementation issues at the expense of high level analysis concerning the requirements; piecemeal modeling by „literal translation" of nouns to entities and verbs to relationships resulting in sub-optimization in the design solution; paying more attention to the syntax of the application than to its semantics; lack of specificity in the universe of discourse.

StoryNet embodies an analysis and design methodology akin to scenario-based design, i.e., the Story Telling Theory (STT) [Faro & Giordano, 97b]. Such methodology provides a template to formalize the use-cases of the Information System in a set of stories and episodes that inherently take into account the user point of view, and allows the designer to organize the requirements in such a way that propagation of effects among stories tend to be confined, thus allowing partial and incremental verification of the model. The method helps lessen the cognitive load involved in performing such task, and provides a contextualized framework that facilitates interpretation of the relevant data and process models.

Thus the prime organizing theme of StoryNet are „stories" and „episodes" that model the organization following the STT design methodology. The template ensures fine-grained indexicality of the materials to identify parts of any project that could be relevant to the domain selected by the student. Attached to these stories and episodes are multimedia documents illustrating the data models and snapshots of the user-interface of the implemented prototype. By examining cases organized in stories and episodes, their structure can be readily perceived, but not the motivations underlying that structuring. This might contribute to a stereotyping problem that has been observed when the students of the IS analysis and design course have access to previous projects but don't have ways to evaluate how they stand relatively to each other. Although methodologically correct, the approach to the analysis tends to re-propose surface level analogies with other organizations, rather than focusing on the differences, as required by truly user-centered design.

To counteract this tendency, cases in StoryNet are incrementally linked in a network of old and new related design cases to highlight how design solutions evolve across time and different groups of learners. Reference links point to units (stories or episodes) that have been used in the design of the retrieved unit and are typed to make apparent how and why that unit was taken into consideration. Available typed links are: 'Correct', 'Extend', 'Detail', 'Adapt', 'Use as is', 'Restructure', 'Other'. Links are mediated by nodes that further qualify the typed conceptual link, by including an additional comment or „explanation". The annotated references of the project should make apparent the analytical and critical effort invested by the student in the project, his or her personal contribution in structuring and introducing new cases, and indirectly foster in the learners greater specificity in their analyses. The structure of the case representation and the network is sketched in [Fig. 1]. StoryNet is currently implemented on Domino Lotus Notes.

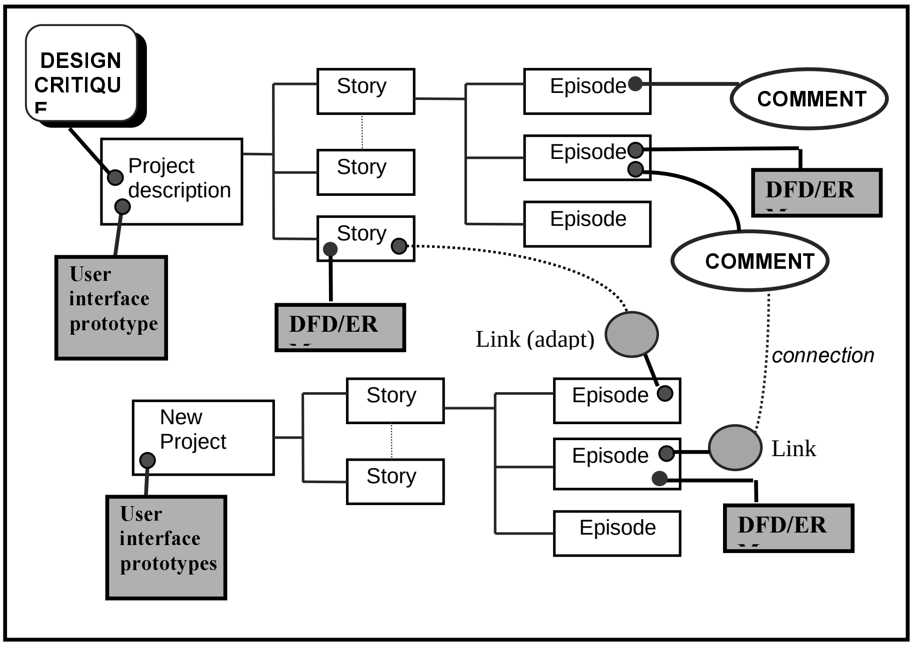

Figure 1: The StoryNet architecture.

An additional component of the shared design memory are the design critiques that are attached to the design representations. Such critiques are based on a guideline for reviewing a project that is an implicit model of how an expert would approach the evaluation of the design. The philosophy of the shared design memory is that there is some added value in sharing such documents beyond the process of producing them. Accordingly, the reason for sharing the critiques is that they have potential to provide a contextualized representation of the possible weaknesses in the reviewed precedents, and function as an intelligible and illustrative warning. The load of creating such resource is distributed among the students and, on average, if each team critique one precedent, as it is required, a representative sample of the biases and misconceptions can be generated.

The relationship between the design features of the shared memory and the hypothesized influence on the biases is sketched in [Fig. 2]

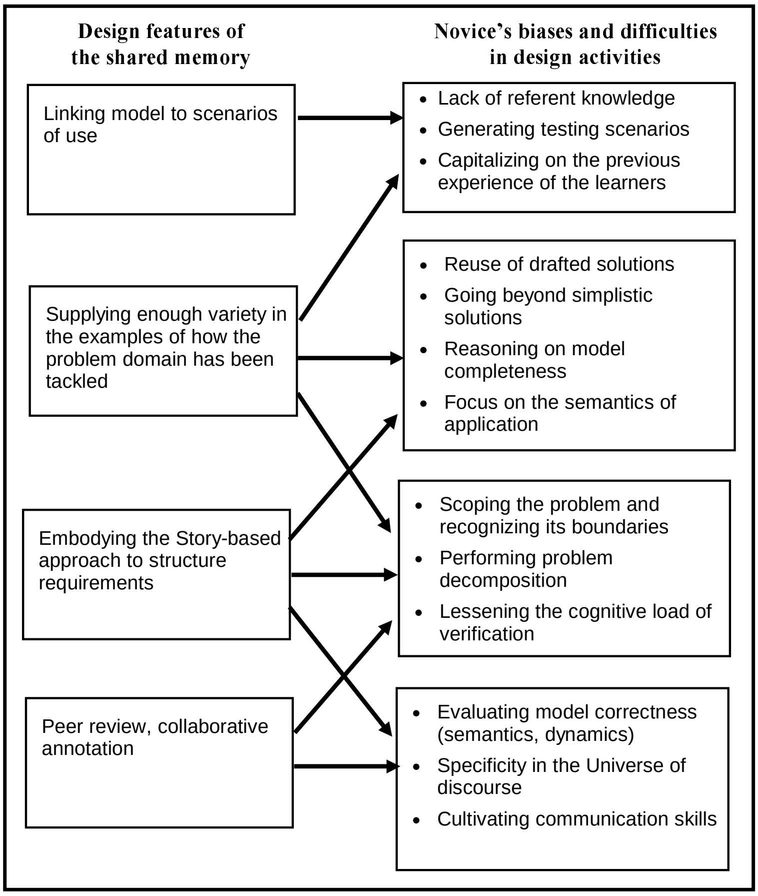

Figure 2: Hypothesized influence of some features of the shared design memory on typical novice biases and difficulties in design activities.

4 The Evaluation Model

As a consequence of the perspective that a shared design memory is just a part of a distributed system, evaluation of its effects on individual and organizational learning must take into account dimensions pertaining to the design artifacts, the community of learners, and the shared design memory, as being an artifact itself. The overall learning environment's effectiveness in supporting analysis and design skills can be evaluated by traditional measures of quality of the design output, although a better picture must include also aspects related to the process of using this environment, such as to what extent students are facilitated in the start-up of their projects, and how the overall process of designing taking into account the shared design memory is perceived.

The underlying model is illustrated in [Fig. 3]. A community of learners comprises small groups and individuals, linked by networks of friendship, trust, recognized expertise Each of these „units“ engages in the activity of producing a design artifact, which is eventually represented in the shared design memory. In the process, they resort also to the precedents available through the shared design memory and possibly contribute some useful knowledge, in the elementary form of a pointer to design resource, or of an articulated design insight.

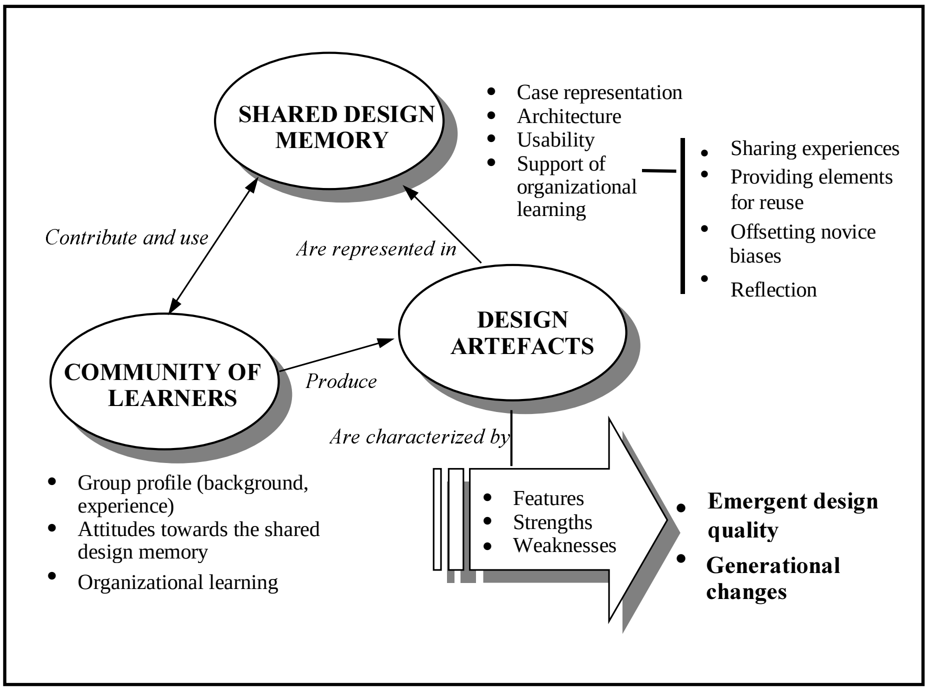

Figure 3: A model to approach the evaluation of a shared design memory.

Relevant dimensions to the community of learners are the group profile, the attitudes towards the shared memory, and the overall organizational learning of the community as a whole. Group profile must be considered when investigating the system because it can be a factor that affects the trends with which the overall system evolves, or that underlies possible failures of specific features of the shared memory to correspond their intended use. Level of initial competence in the design area of interest or in related fields is one aspect to normally include in the definition of the profile. As an heuristic criteria, any features that vary mostly in the overall community (such as schooling background, interest in the subject, to name a few) might be worth attention. Conversely, those features that are strongly defining of the community have to be taken into account too.

Attitudes towards the shared design memory must be seen both globally (in terms of the community's acceptance of the system) and at the individual level (in terms of how each actor responds to the overall mode of dealing and thinking with representations enforced by the system). Thus one angle to investigate - the students' perception of the shared design memory - must regard the pedagogical rationale underlying the approach, i.e., the utility of examining precedents, the utility of critically reviewing some of them and the utility of sharing both the precedents and such reviews. Another angle of analysis must focus on the effort or difficulty involved in the conceptual and communicative operations that are requested of the students in order to sustain and develop the contents of the shared memory. Finally, another angle of analysis must directly address the assumptions made by the designers of the shared memory concerning the effectiveness of the formats chosen to represent the precedents, to link them, and the solutions adopted to browse and to contribute to the system. Related to this latter angle of analysis (and to the first one) are preferences that the users of the system might express on any kind of information that they would like to see in the shared memory.

The organizational learning of the community has to be seen as an emergent effect of micro-activities such as sharing experiences between units that are solving problems collaboratively, reusing design elements, reflecting on the design process, and, overall, as a process of gradual enculturation in the modes of discourse of a professional designer and in the „internal" modes of discourse generated within the community itself. But the key step is relating this global process to the quality of the design artifacts that are produced. Thus one aspect of the evaluation process is to understand it as being relative to the generational changes that occur at each students group turn-over, expressed in the emergent overall design quality that stems from combining the design artifacts features, strengths and weaknesses.

Indicators of the organizational learning that is taking place due to the introduction of the shared memory are:

- the kind of design features or solutions that have become standard (i.e., tend to be present in the majority of the more recent projects, whereas they were rare in a former generation);
- the degree of innovation, i.e., the percentage of projects within a generation that exhibit new features;
- the number and type of design weaknesses, possibly related to the novice cognitive biases and difficulties, that tend to disappear or persist.

To highlight any trend of improvement, quality of designs can be compared to the quality attained in the former years, if the key elements in the used pedagogical approach have been kept constant (except for the introduction of shared memory) and the salient characteristics of the target population of students enrolled in the course for which the shared memory is implemented have not changed. Thus evaluation has an across-generations component and a within-generations component aimed at pointing out any differential effect that might be taking place. In this respect information about the group profile is essential to understand the limits of comparison and generalization.

Concerning the shared design memory, case representation, architecture and usability must be taken into account explicitly because it is necessary to see whether (beyond the raw cases) the quality of the additional representations (e.g., comments, reviews, links) is good enough to be conducive to the hypothesized effect of highlighting from many perspectives design strengths and weaknesses. This amounts to evaluate whether the community is able to produce such statements and to use them productively.

Thus the evaluation method involves a cross-analysis of three types of data originating from: 1) design artifacts; 2) the shared design memory contents, i.e., cases, peer reviews, comments and links; and 3) an individual questionnaire, which collects personal data aimed at characterizing the student population, and asks questions to elicit the student's perception of the effectiveness and usability of the shared memory and a reflection on the kind of difficulties encountered in the design activities.

5 Applying the Evaluation Model: Some Preliminary Results

The first evaluation of StoryNet was conducted based on the data obtained from the period that goes from the initial deployment of the shared memory, March 1997 to the end of July 1997. At this point the shared memory contained 34 new design cases, relevant to the course activity of the then current academic year. At that stage, 56% of the students had taken the final exam (64 students out of 114 enrolled in the course). Data reported on the following originate from the questionnaires collected from the above 64 cases.

5.1 Characterizing the Students Population

The indicators chosen to characterize the student population are: a) the level of experience in database design prior to the course; b) an index of overall competence in skills potentially relevant to the course prior to enrolling at the university; referred to as p.r. (previous related) competence; c) grade point average (g.p.a.); and d) level of interest in the subject. Overall p.r. competence was computed by summing the levels of computer experience, programming experience and database design experience, after asking the students to rate their competence on a five level scale (none = 1, novice = 2, intermediate = 3, advanced = 4, expert = 5).

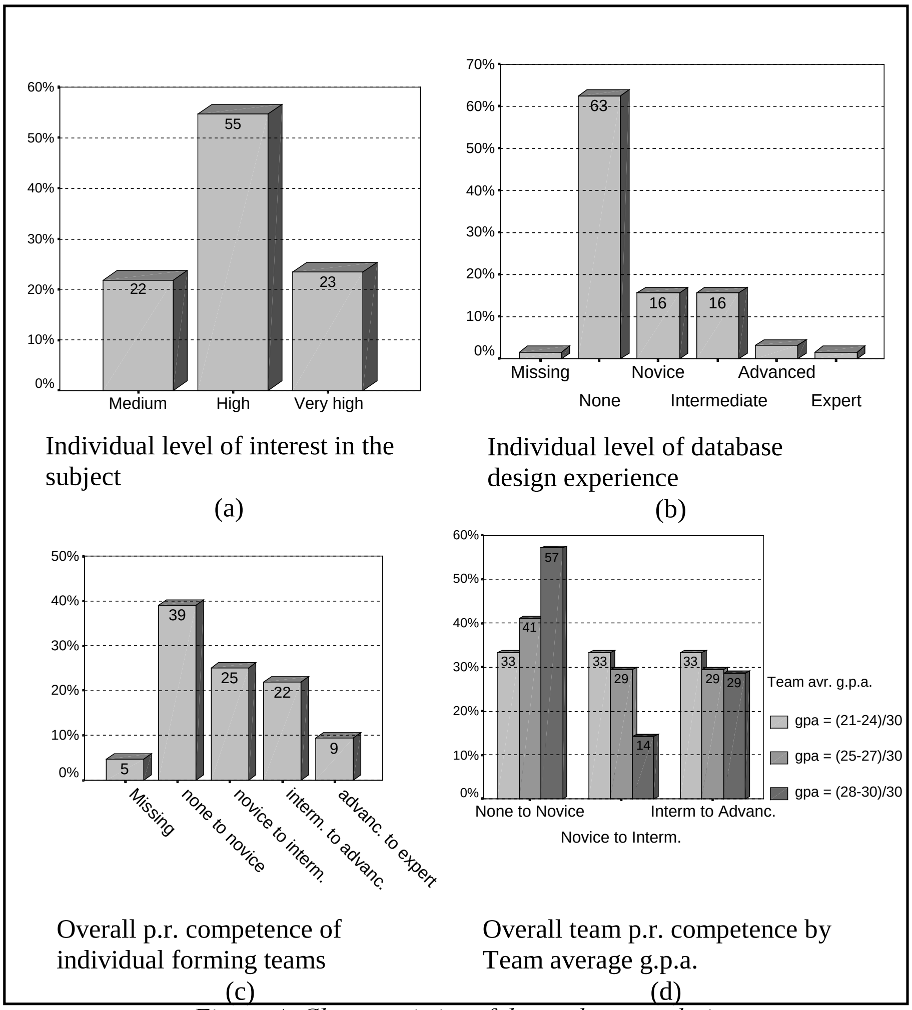

Figure 4: Characteristics of the student population.

The population of students can be characterized as having a strong interest in the subject, with only 22% declaring a medium interest, 58% a high interest and 20% a very high interest [Fig. 4a]; specific prior competence in database design has a majority of cases with none experience (63%), and is skewed towards the none – novice end of the scale, which together account for approximately 80% of the total [Fig.4b].

Of the 34 design projects, 7 were carried out individually, 24 were carried out by teams of two members, and the remaining 3 were carried out by teams of three members. Those individuals who preferred to work alone (11%) did not differ with respect to the mean p.r. competence and mean g.p.a. from those individuals who formed teams.

An analysis of how teams were composed was performed to highlight whether there was any unbalanced distribution of the team characteristics that needed to be considered when evaluating the achieved design quality. The distribution of the overall p.r. competence of individual who formed teams is shown in [Fig. 4c]. There was a tendency of students in the novice to intermediate range of forming groups with students in the intermediate to advanced range, whereas students with none or minimal experience mostly grouped among themselves. An indicator of overall team p.r. competence was obtained by summing the scores of the members and recoding them into three levels. The resulting distribution of the overall team p.r. competence (factored by team average g.p.a.) is shown in [Fig. 4d]. It indicates that the three types of teams had approximately equal proportions of students in the low and medium range of g.p.a., whereas the majority of the students with a high g.p.a. were in the none to novice category of overall team p.r. competence.

There was no correlation between the individual grade point average of the team members, and 58% of the teams were formed by students who had previously worked together.

5.2 Individual Attitudes Towards the Shared Design Memory

The responses concerning the individual perceptions of the shared memory were gathered by an individual questionnaire which ensured confidentiality of the results, and that was administered when the students turned in their projects, before undertaking the oral discussion, and returned at the end of the exam. A total of 64 questionnaires were collected. The first section of [Table 1] shows the ratings of the utility of examining precedents, critiquing them and sharing the critiques; the second section shows the rating of the ease of navigating and contributing to StoryNet.

Data in [Table 1] consistently report perceptions skewed towards the medium-high end of the scale, indicating a substantial agreement on the key tenets of the rationale underlying the use of the StoryNet. There are no particularly revealing differences among frequencies of the responses in the medium category and the sum of the responses in the high-very high categories, except for the utility of sharing the critique, which shows a clear orientation of the responses towards the high-very high categories of the scale, accounting for 54.8% of the cases, versus the 38.7% in the medium category and the 6.4% in the low-very low end of the scale. This result is particularly interesting because, regardless of the perceived utility of actually doing the critique exercise, the students seem to place a very high value on the possibility of looking at the reviews authored by their peers. Interestingly, there is a negative correlation between level of initial design competence and utility of doing the critique exercise (Kendall τ-b = -.281, $p<.05$)

	Rating					m.c.
	very low	low	medium	high	very high	
Utility of examing precedents	-	17.7%	45.2%	32.3%	4.8%	2
Utility of the critique exercise	1.7%	15%	45%	28.3%	10%	4
Utility of sharing the critique	1.6%	4.8%	38.7%	40.3%	14.5%	2
Ease of navigation in StoryNet	4.8%	11.1%	52.4%	27.0%	4.8%	1
Ease of contributing to StoryNet	4.8%	9.55%	49.2%	30.3%	6.3%	1

(*) m.c.= missing cases

Table 1: Individual responses on the tenets of the pedagogical rationale for shared memory and on the overall usability of StoryNet.

For the cases who rated low the utility of precedents (17.7%) the most frequent motivations for holding such view were: poor quality of the available precedents, incompleteness, diversity from the theme to be addressed in the new design case.

Although relevant only to a minority of cases, such responses highlight one risk inherent to using a shared memory constructed by the students themselves: that is, unless specific measures are taken, the weaknesses in the novice artifacts might not be for every student a helpful point to start exercising critiquing skills. However, a mitigating factor for the seriousness of theme diversity as a hindrance, is the start-up effect of the shared memory. In fact, to start the memory, the students were given by the instructor the precedents for review, based on criteria of availability, if relevance to the new design theme could not be satisfied. Therefore, at the time the responses were collected, there was not enough variety of cases formed in StoryNet to provide more relevant examples.

The average number of precedents examined individually is P = 2.36 (SD = 1.39, ranging from 0 to 6) and there is no interaction with level of design experience and level of overall p.r. competence. Concerning the source of the precedents that were examined, 42% of the students resorted to other design cases provided either by colleagues authors of such projects (96%) and/or by colleagues that were not the authors of the design case themselves (30%). These figures are indicative of the social phenomenon of exchanges naturally occurring among the students, and raise the questions of whether and in what respect a shared design memory adds to this process.

The first section of [Table 2] summarizes the agreement level, in a seven points scale, on some statements concerning the conceptual and communicative operations the students were required to perform. The set of questions concerning the links stem from the hypothesis that deploying links entails examining precedents with the aim of finding proper elements for reuse, and that once the design has been carried out, one can still remember what specific aspects of the precedents have exerted influence. This distinction, that might seem unnatural, is due to the fact that in the phase of the start-up of the memory the actual design cases are examined outside StoryNet, and then represented. It is expected that at regimen this splitting will occur to a less extent and that the linking process will support the development and refinement of ontologies in stories and episodes within StoryNet.

The second section of [Table 2] summarizes the agreement level on statements aiming at capturing the possible roles that StoryNet, as a social system and as an instrumental tool, might perform.

The results in [Table 2] show substantial agreement with the statements that:

- to deploy links supports the process of articulating the reasoning that went on during design;
- such links can be useful to peers;
- StoryNet is useful as an index to cases;
- it is a tool that fosters reflection.

This shows acceptance of the system as implemented, and complements acceptance of its pedagogical tenets. On the other hand, the angle of analysis concerning the conceptual operations that are needed to sustain StoryNet reveals a difficulty in the retrospective reasoning that must be performed to articulate what aspects of the current design have been influenced by the precedents (40.6% neutral and 46.3% in the agree side of the scale), and that in the preliminary phase of scanning such precedents it is difficult to anticipate what specific aspects will prove useful in the development of the new project (31.5% neutral and 43.6% in the agree side of the scale). Less difficult is the process of formulating comments in a usable form to peers.

The above difficulties are interesting because whereas there is agreement on the utility of the linking mechanism (response n.1) there is also an indication that it requires an effort that not everybody can sustain, either because of an objective lack of experience that hinders the process of anticipating what might be important, or because there has not been sufficient training in the process of examining the precedents with an attitude towards comparing and contrasting. On the other hand, the difficulty of retrospectively tracing what features have been imported in the new design can also be symptomatic of a process of personal internalization and restructuring of the contents and representational formats that have been encountered. These considerations call for a more detailed analysis concerning whether the process of examining precedents is done more with the spirit of learning or with the spirit of reuse.

	Rating							m.c.
	strongly disagree	disagre e	somewhat disagree	neutral	somewhat agree	agree	strongly agree	
Links express design reasoning	1.6%	-	3.2%	32.3%	25.8%	22.6	14.5%	2
Links are useful to peers	-	4.8%	1.6%	19.4%	37.1%	24.2%	12.9%	2
Links are technically difficult to deploy	15.1%	22.6%	13.2%	18.9%	13.2%	11.3%	5.7%	11
It is difficult to foresee influences of precedents	1.9%	3.7%	7.4%	40.7%	29.6%	5.6%	11.1%	10
It is difficult to isolate influence of precedents	11.9%	1.9%	11.1%	31.5%	25.9%	14.8%	3.7%	10
It is difficult to express comments to be useful to peers	15.9%	14.3%	22.2%	17.5%	9.5%	12.7%	7.9%	1
Cases in StoryNet should be anonymous	39.7%	11.0%	3.2%	30.2%	3.2%	1.6%	3.2%	1
StoryNet is more useful if authors reply to critiques	5.6%	3.7%	9.3%	16.7%	22.2%	13.0%	29.6%	10
StoryNet is useful as an index to cases	-	3.2%	1.6%	25.4%	17.5%	31.7%	20.6%	1
StoryNet can be useful without visioning the whole design case	19.0%	12.7%	11.1%	6.3%	20.6%	19.0%	11.1%	1
StoryNet is useful as a reflection aid anyway	-	1.6%	3.2%	9.5%	30.2%	22.2%	33.3%	1

(*) m.c.= missing cases

Table 2. Responses on the process of links deployment and on the possible roles performed by StoryNet.

There is also agreement on the idea that the StoryNet could benefit if authors respond to the peer reviews. Although StoryNet supports this mechanism, this is unlikely to happen because of the marked asynchronicity with which the groups carry out and complete the design. Indeed, most of the contributions occur just when the project has been completed, shortly before presenting the exam. So any after-course volunteering of information is completely discretional. The deployment of links (from the technical point of view) was rated as being somewhat difficult by approximately 30% of the respondents. This finding had a counterpart: a closer examination of the links actually deployed showed that in some cases the students created the typed links or the explanatory nodes attached to the design cases to be eventually connected but failed to complete the process by skipping the step of actually „closing“ the connection among such nodes. Thus the next version of StoryNet must ameliorate this latent usability problem. On the other hand, the overall ratings of the ease of navigating and contributing to StoryNet are quite satisfactory, as shown in [Table1].

[Table 3] summarizes the opinions about examining in StoryNet different kinds of design representations and the various contributions coming from the participating social actors. Ratings vary on a three points scale: no utility, little utility, and much utility.

	Rating			
	No utility	Little utility	Much utility	m.c.
Stories	-	27.9%	72.1%	3
Episodes	-	32.8%	67.2%	3
Data Flow Diagrams	9.7%	12.9%	77.4%	2
Entity Relation Model	5%	16.7%	78.3%	4
Peer Comments	6.8%	32.2%	61.0%	4
Instructor Comments	3.3%	18.0%	78.7%	3
Links	6.7%	56.7%	36.7%	4
Critiques	6.7%	21.7%	71.7%	4

(*) m.c.= missing cases

Table 3: Relative perceived utility of the StoryNet representations.

A marked diversity of opinion characterizes the issue whether StoryNet can be useful if the whole design case cannot be accessed, which directly brings to bear on the issue of adequacy of the case representational format. Whereas [Table 3] indicates that the supported representations are rated mostly as having much utility, responses to an open-ended section asking whether some other kind on information was needed in the StoryNet indicated the following: a) entire analysis, 2) the software prototype, 3) the dimensioning of the system, 4) other diagrams, 5) best solutions, 6) the instructor evaluation. Only 36% offered this information. [Table 3] shows that the rating of the links is markedly different form the other items, indicating a much more cautious attitude (56.7%) towards the utility of examining links. This result apparently clashes with the responses in Table 1, in which links are rated as a useful vehicle to express design reasoning. A plausible interpretation at this start-up stage, is that for the first generation of students who deployed the first links, the number of links to actually examine in the system was too low to generate the perception that the mechanism is worthwhile representing and has utility as an additional indexing and navigational mechanism.

5.3 Content Analysis of the Critique Exercise

The reviews produced by the teams were analyzed by classifying each critique statement as belonging to one of the following problem areas:

(a) *Dynamics:* concerns the dynamical and simulation aspects of the domain analyzed, and the correct use of the appropriate modeling techniques;

(b) *Breakdowns*: concerns the anticipation of possible breakdowns in the normal functioning of the system and devising exception handling procedures;

(c) *Problem scoping*: concerns the definition of the problem boundaries;

(d) *Structuring*: concerns issues of decomposition and organization, and includes the appropriate use of generalized structures;

(e) *Language*: concerns the specificity or ambiguity of the language used;

(f) *Domain accuracy*: concerns the accuracy of the description of the domain, including the omission of major defining features;

(g) *Design discourse*: concerns all the aspects of data modeling and methodology application;

(h) *Communication*: concerns all the aspects involved in the effective communication of the analysis and design representation, from the graphical layout to the explanatory comments to the design representations.

A total of 32 reviews were analyzed, yielding the distribution of detected weaknesses shown in [Fig. 5].

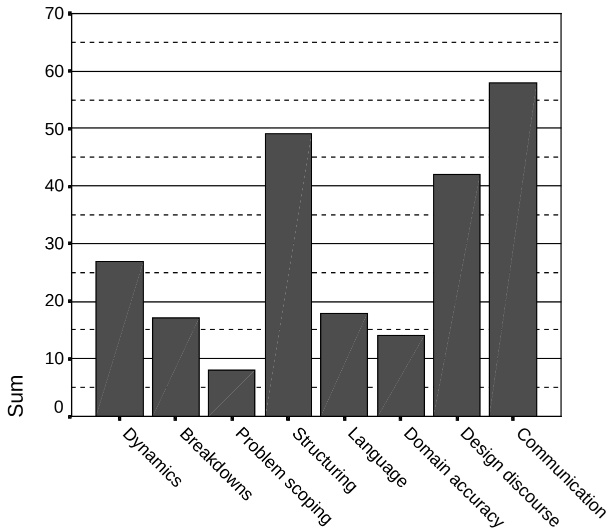

Figure 5: Distribution of the type of weaknesses highlighted in the reviews of precedents.

For the purposes of the present analysis, it can be remarked that the distribution shown in the graph must not be interpreted as representing the average quality of the reviewed precedents. Rather it is a representation of what was perceived by the students at their current level of maturation, given the „observation lens“ of the critique guideline, the particular precedent examined, and their expressive ability. In fact, the most frequent counts in a problem area could be accounted for partly by a better sensibility of the reviewers towards the underlying issues. Thus the distribution is better interpreted as meaning that the overall system, socially, was able to find instances of issues to be aware of in the whole spectrum of problem areas, certainly with a varying degree of penetrating insights, but demonstrating that there was no „organizational blindness“ to any if them.

Each review was also coded as to whether it was effective in highlighting the strengths and weaknesses of the precedent, by meeting the criteria of clarity of expression and localization, i.e., justification by direct reference to an easily identifiable part of the precedent itself. Additionally, a statement was considered as: a) conveying domain knowledge when the authors justified their evaluation making explicit reference to the application domain, and b) conveying design knowledge when the authors justified their evaluation by proposing design solutions, or by introducing a novel dimension other than those suggested in the review guidelines.

Approximately, half of the contributors were able to communicate effectively the strengths and weaknesses of the reviewed cases, and approximately 40% did so by referring to the domain application.

A similar pattern occurs for the links among the design cases. Of the overall contributing teams or individuals, 85% placed links to the precedents. The average number of links placed was 2 (ranging from 1 to a maximum of 4). The majority of the links were of the type „Adapt“ (24%), „Use partially“ (18%), and „Other“ (18%).

However, the links that cogently expressed design reasoning beyond the synthetic type were only 32% of the total. Relatively, precedents' reviews appeared to be a medium more easily tapped to augment the knowledge in the shared memory.

The findings that 40% of the reviews and approximately 30% of the links explicitly convey knowledge are important to estimate the actual potential of the shared memory for highlighting design weaknesses and strengths in context (when the number of cases increases) and the information gain with respect to the knowledge implicitly conveyed by the case representation itself.

6 Design Changes across Generations: Evaluating Design Quality

Although there are some „intrinsic“ qualities of the conceptual model as a representation (such as its syntactic correctness or its readability) these are second order regard to the issue that the semantic of the model is correct and complete with respect to the domain and the activities that the information system is supposed to support. *Syntactic* quality, *semantic* quality, and *pragmatic* quality (i.e., that the model is understood by the interested parties) must be addressed in terms of feasibility goals [Lindland et al., 94]. Total validity and completeness, defined, respectively, as the property that all statements made by the model are correct and relevant to the problem, and as the property that the model contains all the statements about the domain that are correct and relevant, cannot be achieved except for very simple problems. The goals of semantic quality are *feasible validity* and *feasible completeness* of the model, and the goals of pragmatic quality are *feasible comprehension* of the model. Quality properties of the model such as correct, minimal, annotated and traceable, consistent, unambiguous are mostly subsumed by validity and completeness. Feasibility provides a stop rule to terminate the modeling activity, i.e., when the model has reached a state where further modeling is less beneficial than applying the model in its current state.

Quality of information is related to how data are used, thus quality of the data generated by the information system should not be couched in terms of „data centric“ concepts such as entities, attributes and values, but in relation to intrinsic dimensions, assuming that the true intentions of the users have been captured in the conceptual model [Wand & Wang, 96]. Such intrinsic dimensions are: 1) *Complete* (there is no loss of information about the application domain, 2) *Unambiguous* (data generated by the information system cannot be interpreted in more than one way, 3) *Meaningful* (data can always be interpreted in a meaningful way, and 4) *Correct* (data derived from the information system conform to those used to create these data. This view complements the notion of feasible quality in the conceptual model by emphasizing the mapping process from requirements to data, in such a way to project the data model towards aspects of design more oriented the prototyped information system.

An important overlooked dimension in traditional approaches to quality is how the conceptual and data model that are being crafted handle issues of *interactivity* of the overall system, also considering that within a scenario-based approach to design (such as STT) the interface of the system is not a separate aspect of the design. A way to indirectly address this aspect of quality is to place emphasis also on the modeling of the dynamic interactions in the system, resorting to different representations to carry through the design the user and his or her interactions with the data. This kind of modeling effort is also beneficial for the consistency of the data.

6.1 A Framework for Evaluating Design Quality

The framework adopted for evaluating the design quality is centered on the following points:

- Semantic quality of the overall conceptual model, especially focused on feasible validity;
- Emphasis on completeness meant as multidimensionality of analysis relative to the stories/episodes identified by the students, which set the internal standard for evaluating the design;
- Feasible completeness and validity of data modeling, relaxed with respect to the fact that small errors cease to assume relevance if they are corrected or disambiguated in further representations, often complemented by the prototype;
- Integration of a dimension that explicitly takes into account interactivity;
- Communication, to address feasible pragmatic quality of the design.

Accordingly, the design scoring template has been developed following a design features approach (i.e., grading 0 or 1 - either the absence or presence of a certain design feature, mostly concerning representational or stylistic solutions) and a graded quality evaluation for Domain modeling specificity and Data modeling (0.5 if a minor weakness is present, 0 if a major weakness is present and 1 for feasible validity).

The feature-based approach is particularly appropriate in our case because whereas in other studies the designs are comparable because they deal with the same problem, in this context it is necessary to find a common reference point across designs that tackle different problems. Also, a feature-based analysis is a more effective way to detect, from an organizational learning point of view, what tends to become consolidated, what is emerging or tends to disappear generation after generation.

The set of features and their values have been mapped to contribute to the definition of the values of the following quality indicators:

- *Domain modeling specificity*
- *Structuring (decomposition)*
- *Structuring (composition)*
- *Data modeling*
- *Dynamics of interactions*
- *Breakdown*
- *Communication*
    - *textual communication*
    - *visual/hypertextual communication*

The list of features that concur to form the indicators and the range of variation of the scores is reported in [Table 4].

Domain modeling specificity	**Structuring (Decomposition)**	**Structuring (Composition)**
range : 0-4	range: 0-8	range : 0-6
• Ratio of domain specific entities to total entities • Ratio of domain specific functions to total functions • Ratio of domain specific interactions to total interactions • Use of organization charts	• Use of „consist of“ structures • ELH associated to the episodes • PO associated to the episodes • DFD structured by episodes • ERM structured by episodes • Prevalence of representations of actors vs. functions • Use of story/entity matrix • Top-down decomposition of episodes	• Use of colored arcs in the DFD to identify stories/episodes • Use of colored functions in the DFD to identify stories/episodes • Link from ELH to episodes • Link from ERM to episodes • Diagrams obtained as incremental block composition • Top-down DFD within episodes
Data modeling	**Dynamics of interactions**	**Breakdown**
range : 0-11	range : 0-6	range : 0-2
• Analysis by aspects • Use of „consist of“ structures • Datastores correctness • From entities to relational tables transformation • Relational tables partitioned by stories • ERM correctness • Relationships correctness • Use of „is a“ structures • Ternary relationships • Normalization • Normalization correctness	• Animated ELH • Episodes tree • Animated episode tree • DFD animated in synchronization with the episode • Animated Petri networks • Petri Networks with animated captions	• Insightful What Can Go Wrong • What Can Go Wrong represented in the episode tree

Legend: DFD = Data Flow Diagram; ERM = Entity Relation Model; ELH = Entity Life History; PO = Process outline; PN = Petri Network

Table 4: Features in design quality indicators and range of variation of the scores.

Communication (textual)	**Communication (visual-hypertextual)**	**Index of design Complexity**
range : 0-4	range : 0-13	
• Commented matrices • Introduction to the methodological approach • Introduction to the domain • Description of high level scenarios and use cases • Commented organization charts	• Iconic diagram boxes • Iconic arcs • Use of colored arcs in the DFD to identify stories/episodes • Use of colored functions in the DFD to identify stories/episodes • Loose animation: PN presented in blocks • animated entities in ERM • „Live“ entities • „Live“ relationships • „Live“ arcs • Animated normalization • Petri Nets with animated captions • Animated Petri nets • DFD animated in synchronization with the episode	Computed as the sum of the following : • N. of stories • N. of episodes • N. of functions • N. of entities • N. of interactions in the DFD • N. of entities and relationships in the global ERM • max (states and arcs in the global ERM) • max (transition, places in the Petri Nets) • max (arcs and states in the automata)

Legend: DFD = Data Flow Diagram; ERM = Entity Relation Model; ELH = Entity Life History; PO = Process outline; PN = Petri Network

Table 4 (cont.): Features in design quality indicators and range of variation of the scores.

The novice difficulties and biases mentioned in [Section 3] are mapped in the above dimensions as follows. „*Domain modeling specificity*“ is an indicator the specificity of the language, derived from the ratios of domain specific entities to total entities, domain specific functions to total functions, and domain specific interactions to total interactions. „*Breakdowns*“ addresses depth of analysis and indirectly indicates the ability to generate testing scenarios against which evaluating the quality of the data model. „*Structuring (decomposition)*“ aggregates the design features that indicate an effort toward organizing the decomposition the static and dynamic data models by one or more of the following strategies: 1) decomposition of the static and dynamic data models following the story/episode approach, or the top-down approach, or both; 2) use of „consist of“ structures in modeling the entities; 3) use of cross-reference matrices between stories and entities. „*Structuring (composition)*“ aggregates the design features that indicate an effort towards maintaining in a global representation explicit information about the elements that are aggregated and composed, either by color-coding or with links, or by simulating through animation the incremental creation of the global representation.

6.2 Results

To evaluate the trend of design quality a random sample of 16 projects was drawn from the pool of 32 cases belonging to the new generation (henceforth, 2nd generation) and a matching set was created by selecting the corresponding reviewed projects from the old generation (henceforth, 1st generation). Because one of the reviewed precedents was missing and two cases from the sampled 2nd generation had reviewed the same precedent, the matching set was completed by adding two other cases drawn from the pool of precedents available to the students. The projects were scored blindly by an independent rater, an expert in the field, following the template elaborated in the previous section. Data from the two generations were screened separately for outliers and plotted for a qualitative analysis of variations in their distributions. A t-test was performed to compare the mean score of each quality dimension for the two generations of design cases. A correlational analysis using the Spearman ρ rank correlation coefficient was performed to investigate the relationships among the various dimensions of the design quality for the second generation, taking into account the characteristics of the teams. Also, a correlational analysis for elementary design features and generations was performed to highlight, at a micro level, what were the practices, or representational modes starting to become diffused in the new generation. The descriptive statistics for the dimensions of design quality are reported in [Table 5].

		Descriptives					
	Gen.	Mean	Median	Std. Deviation	Minimum	Maximum	Range
Domain modelling specificity	1st Gen.	2.63	3.00	1.02	.00	4.00	4.00
	2nd Gen.	3.40	3.00	.79	1.00	4.00	3.00
Structuring (Decomposition)	1st Gen.	2.75	2.00	1.45	1.00	6.00	5.00
	2nd Gen.	3.97	4.00	2.16	1.00	7.00	6.00
Structuring (Composition)	1st Gen.	2.19	2.00	1.52	.00	6.00	6.00
	2nd Gen.	2.00	2.00	1.15	.00	4.00	4.00
Data Modelling	1st Gen.	5.28	5.00	1.61	1.50	8.00	6.50
	2nd Gen.	5.34	5.25	2.58	.00	9.00	9.00
Dynamics of interactions	1st Gen.	1.06	1.00	.57	.00	2.00	2.00
	2nd Gen.	2.25	2.00	1.06	.00	5.00	5.00
Breakdown	1st Gen.	1.53	1.75	.56	.50	2.00	1.50
	2nd Gen.	1.38	1.75	.74	.00	2.00	2.00
Textual communication	1st Gen.	1.44	1.00	.63	1.00	3.00	2.00
	2nd Gen.	2.44	2.50	.79	1.00	4.00	3.00
Hypertextual communication	1st Gen.	1.94	2.00	1.22	.00	4.00	4.00
	2nd Gen.	3.63	3.50	1.71	1.00	7.00	6.00
Complexity	1st Gen.	121.56	107.00	36.01	79.00	201.00	122.00
	2nd Gen.	121.41	123.50	34.02	61.00	185.00	124.00

(n. of cases: 1st Gen = 16, 2nd Gen. =16)

Table 5: Descriptive statistics for design quality dimensions in the 1st and the 2nd generations.

Additional information about the trends in the two generations can be derived from the concise representation of the distribution of the scores for each variable, as shown in the boxplots in [Fig. 6]. The shaded area in the boxplot represents the range within which fall 50% of the scores (25^{th} to 75^{th} percentile) and the line in the box indicates the median. The vertical lines extending out of the box mark the largest and the smallest observed value that is not an outlier, i.e., more that 1.5 box-lenghts from the upper and lower edge of the box. The distribution of *Domain modelling specificity* in the 1^{st} generation (with 12 cases who scored 3, three cases who scored 0 or 1 and only one case who scored 4) reflects the bias of lack of specificity or „stereotyping“ frequently observed in the precedents. The situation is significantly better in the 2^{nd} generation, in which the lowest value is 2.50 and 15 out of 16 cases fall between 3 and 4. *Structuring (Decomposition)* reveals the significant effect for the 2^{nd} generation regarding the increase in the structuredness of design. This indicates that some of the features, or different combination of features relative to modes of organizing and layering the design are becoming common. *Structuring (composition)* did not show any change across the generations. *Data modeling* reveals that for the 2^{nd} generation there was an increase in the spread of the values, and an increase in the number of cases in the higher range if the scale, thus indicating a tendency to improve. For the variable *Breakdown* there was no substantial variation, only the fluctuation of few cases in the lower end of the scale. The index of complexity remained practically unchanged across the two generations.

Statistically reliable differences among the two generations were found for: *Domain modeling specificity* ($t= -2.68$, $p<.05$); *Structuring (decomposition)* ($t= -2.12$, $p<.05$, $df = 25$ - adjusted for inequality of variance in the samples -Levene test), *Dynamics of interactions* ($t= -3.92$, $p<.05$), *Communication (textual)* ($t= -3.95$, $p<.05$), and *Communication (visual-hypertextual)* ($t= -3.21$, $p<.05$).

The above results indicate that the projects of the 2^{nd} generation showed more specificity in the language used to describe the domain, more effort aimed at structuring the analysis and an increased attention towards modeling the dynamics of interactions. Tapping more communication modalities, both textual and hypertextual, was also a distinctive mark of the 2^{nd} generation.

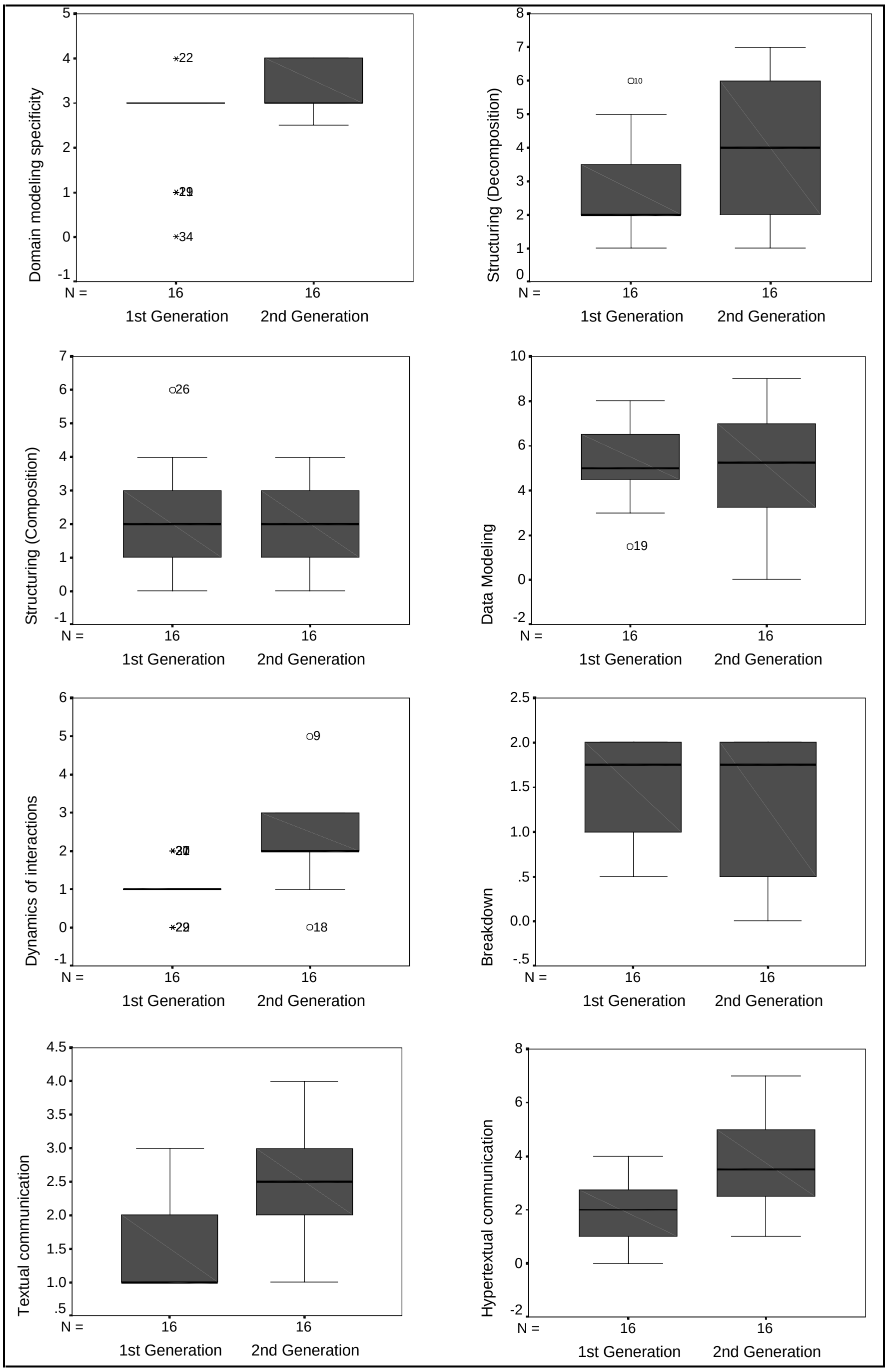

Figure 6: Qualitative distributions for the design quality variables.

Correlations

		Domain modeling specificity	Structuring (Decomposition)	Structuring (Composition)	Data Modelling	Dynamics of interactions	Breakdown	Textual communic.	Hypertextual communic.	Team overall p.r. competence	Team average g.p.a.
Spearman's rho	Domain modeling specificity	1.000	-.433	.313	-.476	-.154	-.213	.263	-.341	.092	-.207
	Structuring (Decomposition)	-.433	1.000	.428	.867**	.306	-.047	.101	.538*	.415	.425
	Structuring (Composition)	.313	.428	1.000	.462	.074	-.011	.300	.292	.339	.355
	Data Modelling	-.476	.867**	.462	1.000	.076	-.060	.035	.456	.316	.562*
	Dynamics of interactions	-.154	.306	.074	.076	1.000	.421	.000	.761**	.231	-.163
	Breakdown	-.213	-.047	-.011	-.060	.421	1.000	.288	.102	-.213	-.262
	Textual communication	.263	.101	.300	.035	.000	.288	1.000	-.113	.113	-.042
	Hypertextual communication	-.341	.538*	.292	.456	.761**	.102	-.113	1.000	.182	.038
	Team overall p.r. competence	.092	.415	.339	.316	.231	-.213	.113	.182	1.000	-.055
	Team average g.p.a.	-.207	.425	.355	.562*	-.163	-.262	-.042	.038	-.055	1.000

**. Correlation is significant at the .01 level (2-tailed).

*. Correlation is significant at the .05 level (2-tailed).

Table 7: Correlations among the dimensions of design quality and team characteristics for the sample of the 2nd generation.

The analysis of the correlations among the design dimensions for the 16 cases of the 2nd generation highlights a set of relationships, shown in [Table 7], that can be helpful in interpreting the effects achieved with the shared memory. The following set of correlations is of interest:

1) *Data modeling* is correlated with *Structuring-decomposition* (Spearman $\rho = .867$, $p<.05$); *Structuring- decomposition* and *Communication visual-hypertextual* are correlated also (Spearman $\rho = .538$, $p<.05$);

2) *Dynamics of interactions* is correlated with *Communication visual-hypertextual* (Spearman $\rho = .761$, $p<.05$).

Although the development of a model is beyond the scope of the paper, these two broad partitions suggest viewing design quality as emerging from the interplay of two fundamental activities, that occur at the outset of the design process, i.e., structuring and finding ways to communicate that incorporate visual, temporal, and hypertextual modalities. These reverberate on the quality in data modeling and the quality in the analysis of the dynamics of interactions.

The correlations suggest that:

- Increased structuring co-occurs with increased quality in the data modeling;
- Increased hypertextual communication, which inherently carries a temporal component (related to animations) and a structuring component (related to the layering of information), co-occurs with a better understanding and quality of the dynamical aspects of the system being designed.

This latter result is consistent with the results of the qualitative analysis of the generations, because generational increases in hypertextual communication and structuring co-occurred with the generational increases in dynamics of interactions.

However, the significant increase in structuring did not result in an equivalent increase for the data modeling dimension (which only improved slightly across the two generations). The positive correlation between *Data modeling* and *Team g.p.a.* (Spearman $\rho = .597$, $p<.05$). might explain why an aspect such as this is more resistant to be self-corrected by the effect of the shared memory. A higher g. p.a., at this stage in the engineering studies (second year) is strongly indicative of general analytical abilities, and this could be a factor that accounts for more variance in the quality achievable in the data modeling activities than factors such as more sophisticated ways of tackling decomposition, increased exposure to cases, o some degree of experience with database design.

Interestingly, the dimensions that significantly improved in the 2^{nd} generation (i.e., specificity of the language, dynamics of interactions and quality of textual and hypertextual communication, all equally fundamental for the overall pragmatic quality of the design) were not correlated with the design team characteristics [Table 7]. The various types of teams were all represented in the sample. This suggests, conservatively, that in the conditions created by StoryNet, team typology does not determine any particular advantage or disadvantage concerning the possibility to benefit from StoryNet, at least in those dimensions susceptible to improvement. This is quite encouraging, given the widespread variety in the type of team composition [Fig. 4d] that is inherent in the social context in which StoryNet is functioning.

A feature by feature analysis of the projects reveals some elements that distinguish the two generations, and indicates, at the level of features, what are practices or representational modes that are becoming diffused in the new generation [Fig. 7]. For example, the use of animated Petri Networks was quite rare in the old generation, but is a widespread feature in the new generation; less dramatic, although significant, is the effect for the additional self-explanation provided by animated captions. More diffused are also organization chart-like representations to give more concreteness and structure the domain description, the use of commented cross-reference matrices and an overall approach to structuring that is more centered around „actors“ rather than functions. Also, it is interesting to note the appearance of a feature that was presumably absent or very infrequent in the old generation, i.e., the animation of the interactions represented in the DFD in synchronization with the unfolding of the episode in which they occur.

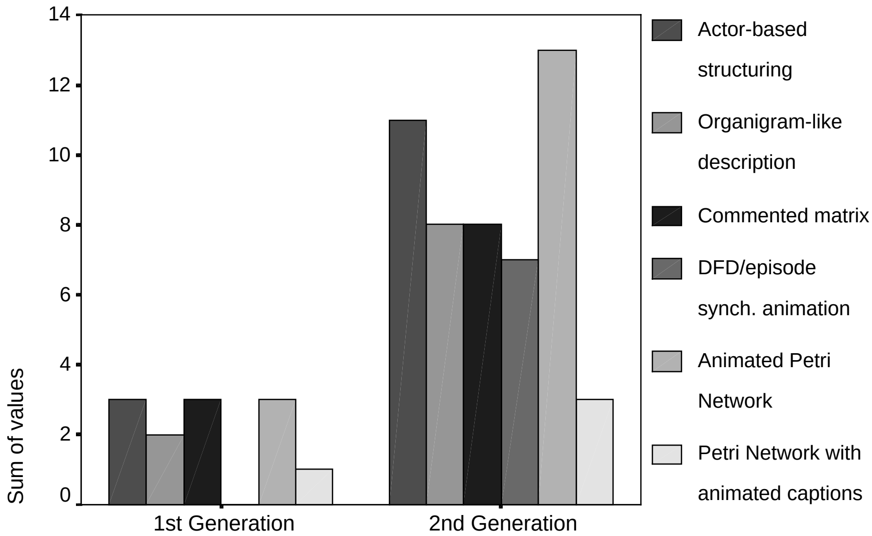

Figure 7: Design features that appear or become more diffused in the „new“ generation.

6.3 Discussion

The most remarkable effect highlighted by the evaluation is that the community of learners, supported by the StoryNet shared memory, has been able to orient itself towards a way of approaching design that has led to measurable improvement in key dimensions of design quality.

Such improvement can be better described as achieving more balanced designs, in which the data modeling component that was at the core of the previous designs has been complemented and enriched by a component addressing all the dynamics of interactions, at the process and the event level, in a way strongly tied to story-based structuring, and oriented towards impacting in the interactions in the prototype. This emergent quality seems to be sustained by the emphasis that the learners have placed on structuring and communicating according to visual, temporal, and hypertextual modalities, and on explaining and justifying key analysis and design steps. This has led to a better semantic and pragmatic quality of their models. Also, structuring and communication (and their impact on the dynamics of interaction) are particularly important because the students are dealing with the analysis and design of user-centered systems that have a strong interactive component.

The fairly close match between the improvements and the profile of the weaknesses highlighted in the design reviews [Fig. 5] suggests that applying the review guidelines (and personally finding concrete instances that made them meaningful) contributed to perceiving structuring, dynamics and communication as very important. If structuring is „internalized“ because many examples of how to structure stories and episodes are available (and thus the relevant bias on decomposition is offset) on the other hand, the reverberating effect of StoryNet on quality of communication can be triggered also by one critical example, for example having to review one precedents that is very hard to comprehend.

Also, a micro-process of diffusion of design features across generations was observed. At the individual, concrete level, this process probably proceeds by incorporating, initiating or transforming small design features or representational devices that the students like or perceive to be supportive of the aspects they want to improve. Interestingly, the features belong to the communication –visual hypertextual and structuring dimensions. By being „imitated“ they have become part of the culture, part of what it is expected to be found in a good design, and contributed to the organizational learning of the community. What takes place in the community it is not simply „reuse“ of scenarios or pieces of software, but it is also reuse and interpretation of representational devices.

Thus the mode of discourse promoted by StoryNet is not only critical evaluation of models, reuse and adaptation of components (consistent with authentic activities in the IS domain), but it is also finding ways to become attuned to what is deemed important in the community, and participating personally in the re-definition of such perception.

Because of the start-up phase of the StoryNet, it is too early to draw conclusions concerning the reasons why other dimensions did not improve as much, for example, the quality of the breakdown analysis, or the overall data modeling. However, because these aspects can be related to the diffusion of domain knowledge and to the generation of memorable examples (which depend on the growth of shared memory) the preliminary evaluation already suggests an avenue for action. Given that approximately only 30% –40% of the annotated knowledge in the links and the reviews reaches standards of clarity and effectiveness some effort must be placed in introducing more sophisticated search and visualization mechanisms, to quickly address the documents that enjoy such quality and facilitate the process of diffusing them as a reference resource.

7 Concluding Remarks

The observed improvement can be ascribed to the activation of the StoryNet system on different grounds. First, all the other elements of the pedagogical setting (the program, the professor responsible for the course, the labs) have not changed in the last two years, except for the introduction of StoryNet and related activities. Second, a change in the characteristics of the population of students who enroll in the course is very unlikely. In any case the preliminary results show that even teams with markedly different characteristics benefit (although may be for different reasons) from StoryNet, and the whole range of teams typologies was represented in the sample. Third, the informal practice of exchanging precedents was already established, but without making this process „formal“ and „official“, and framing it in the modality of reviewing and acknowledging precedents influences and reuse, the overall quality achieved in the former generation remained confined in the usual boundaries. By formalizing this process, more people could access more precious resources (that otherwise would have remained outside their reach), but, most importantly, it was formalized in a way that was acceptable to its users.

The results of the analysis of the data provide an additional perspective in the learning of the novice designers that would have not been emerged so clearly by relying only on participant observation, although the model for approaching the evaluation stems from one year of direct involvement in the process of designing StoryNet and interacting with the students. The opportunities that have been actualized are: an increase in the quality standards for the designs; knowledge that starts being represented in StoryNet, in addition to the „raw“ cases; and seeds for new improvements that are becoming available to the community. Also there are some concrete indications as to what can be done to improve the process of shared memory (i.e., link deployment and more advanced search and visualization techniques) and an indicative estimate of the leverage that can be expected. We need finding ways to strike the notes that allow the community of learners to perceive and act on aspects that so far were underplayed (e.g., design breakdown anticipation and data modeling optimization), without giving up the achieved benefits. In fact, it is not obvious or to be taken for granted that the achieved balanced quality of the designs will be sustained across the coming generations, although there are good reasons to expect it will. Thus it is important that the evaluation process be continuously kept in the background. Further research can address the more specific dynamics of evolution of the shared memory and the role played by each actor or type of team in it.

References

[Ackermann & Malone, 90] Ackermann, M. & Malone, T.: „Answer Garden: a Tool for Growing Organizational Memory“; Proc. Conf. on Information Systems, ACM Press, (1990), 31-39.

[Batra & Antony, 94] Batra, D. & Antony, S. R.: „Novice Errors in Conceptual Database Design“; European Journal of Information Systems, 3, 1(1994), 57-69.

[Barber et al., 92] Barber, J. et al.: „AskJef: Integrating Case-based Reasoning and Multimedia Technologies for Interface Design Support ; in J. Gero (Ed.) Artificial Intelligence in Design; Kluwer Academic Publisher, Boston (1992).

[Brown, 92] Brown, A. L.: „Design Experiments: „Theoretical and Methodological Challenges in Evaluating Complex Intervention in Classroom Settings“; The Journal of the Learning Sciences, 2, 2 (1992), 141-178.

[Brown et al., 89] Brown, J. S., Collins, A., & Duguid, P.: „Situated Cognition and the Culture of Learning“; Educational Researcher, 18, 1 (1989), 32-42.

[Brown & Duguid, 91] Brown, J. S. & Duguid, P.: „Organizational Learning and Communities of Practice: Towards a Unified View of Working, Learning and Innovation“; Organization Science, 2, 1 (1991), 40-57.

[Bruegge & Coyne, 94] Bruegge, B. & Coyne, R. F.: „Teaching Iterative and Collaborative Design: Lessons and Directions“; Proc. 7th SEI CSEE Conference on Software Engineering Education, San Antonio, Texas, USA, January 1994; Springer Verlag, 411-427.

[Buckingham Shum, 97] Buckingham Shum, S.: „Negotiating the Construction and Reconstruction of Organisational Memories“; Journal for Universal Computer Science, 3, 8 (1997), 899-928.

[Conklin & Begeman, 88] Conklin, J. & Begeman, M. L.: „gIBIS : a Hypertext Tool for Exploratory Policy Discussion“; ACM Transactions on Office Information Systems, 6, 4 (1988), 303-331.

[Faro & Giordano, 97a] Faro, A. & Giordano, D.: „StoryNet: an Evolving Network of Cases to Learn Information Systems Design“; in N. Jajaratna (Ed.), Proc. 5th BCS Conference on Information Systems Methodologies; Springer-Verlag, Preston, UK (1997).

[Faro & Giordano, 97b] Faro, A. & Giordano, D.: „Between Narration and Drama: Information Systems Modelling Revisited“; in R. Galliers et al. (Eds.), Proc. of ECIS'97, 5th European Conference on Information Systems; Cork Publishing Limited, Cork, Ireland (1997), 383-395.

[Giordano, 97a] Giordano, D.: „Learning by Building Organisational Memories: Some Issues.“ Proc. EAEEIE 8th Conference, European Association for Electrical and Information Engineering Education; EAEEIE, Edinburgh, Scotland (1997).

[Giordano, 97b] Giordano, D.: "A case-based pedagogy to learn Information Systems design"; Online Proc. 11th Annual Midwest Computer Conference, Springfield, University of Illinois at Springfield (1997)

http://www.uis.edu/%7Emcc97/html/papers/giordano/

[Huber, 91] Huber, G. P.: „Organizational Learning: the Contributing Processes and the Literatures“; Organization Science, 2, 1(1991), 89-115.

[Kolodner, 93] Kolodner, J.: „Case-based Reasoning“; Morgan Kaufmann, San Mateo, CA (1993).

[Koschmann, 96] Koschmann, T. (Ed.): „CSCL: Theory and Practice of an Emerging Paradigm“; Lawrence Erlbaum, Hillsdale, NJ, (1996).

[Koschmann et al. 96] Koshmann, T., Kelson, A. C., Feltovich, P. J., & Barrows, H. S.: „Computer-supported Problem-based Learning: a Principled Approach to the Use of Computers in Collaborative Learning“; in T. Koschmann (Ed.), CSCL: Theory and Practice of an Emerging Paradigm; Lawrence Erlbaum, Hillsdale, NJ, (1996), 83-124.

[Lave, 88] Lave, J.: „ Cognition in practice“; Cambridge University Press, New York (1988).

[Levitt & March, 88] Levitt, B. & March, J.: „Organizational Learning“; Annual Review of Sociology, 14 (1988), 319-340.

[Lindland et al., 94] Lindland, O. I., Sindre, G., & Sølvberg, A.: „Understanding Quality in Conceptual Modeling“; IEEE Software, 11, 2 (1994), 42-49.

[Oshima et al., 96] Oshima, J. Scardamalia, M. & Bereiter, C.: „Collaborative Learning Processes Associated with High and Low Conceptual Progress“; Instructional Science, 24 (1996), 125-155.

[Pea, 93] Pea, R.: „Practices of Distributed Intelligence and Design for Education“; in G. Salomon (Ed.), Distributed Cognitions; Cambridge University Press, New York (1993).

[Salomon, 93] Salomon, G. (Ed.): „Distributed Cognitions“; Cambridge University Press, New York (1993).

[Scardamalia & Bereiter, 93] Scardamalia, M. & Bereiter, C.: „Technologies for Knowledge-building Discourse“; Communications of the ACM, 36, 5 (1993), 37-41.

[Silverman, 95] Silverman, B. G.: „Computer Supported Collaborative Learning“; Computers in Education, 25, 3 (1995), 81-91.

[Spiro et al., 91] Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L.: „Cognitive Flexibility, Constructivism, and Hypertext: Random Access Instruction for Advanced Knowledge Acquisition in Ill-structured Domains“; Educational Technology, 31, 5 (1991), 24-33.

[Sutcliffe & Maiden, 92] Sutcliffe, A. & Maiden, N.: „Analysing the Novice Analyst: Cognitive Models in Software Engineering“; Int. Journ. of Man-machine Studies, 36 (1992), 719-740.

[Walsh & Ungson, 91] Walsh, J. P. & Ungson, G. R.: „Organizational Memory“; Academy of Management Review, 16, 1 (1991), 57-91.

[Wand & Wang, 96] Wand, Y. & Wang, R.Y.:"Anchoring Data Quality Dimensions in Ontological Foundations"; Communications of the ACM, 39, 11 (1996), 86-95.