# INTERNATIONAL JOURNAL OF INSTRUCTIONAL TECHNOLOGY AND DISTANCE LEARNING

**March 2005**

**Volume 2 Number 3**



**ISSN 1550-6908**

## PUBLISHER'S DECLARATION

The International Journal of Instructional Technology and Distance Learning is refereed, global in scope, and focused on research and innovation in teaching and learning.

The Journal was established to facilitate collaboration and communication among researchers, innovators, practitioners, and administrators of education and training programs involving instructional technologies and distance learning. The editors and peer reviewers are committed to publish significant writings of high academic stature.

The initial year of publication was funded by the TEIR Center, Duquesne University. The Executive Director of the Center, Lawrence Tomei, served as Publisher. Additional support was provided by DonEl Learning Inc. and freely donated time of the editors, and peer reviewers.



Donald G. Perrin
Executive Editor

International Journal of
Instructional Technology & Distance Learning

**Vol. 2. No. 3.**
**ISSN 1550-6908**

# Table of Contents – March 2005

| | Page |
|---|---|
| **Editorial: The Future of Learning**<br>Donald G. Perrin | 1 |
| ***Refereed Papers*** | |
| **Normative Specifications of Learning Objects and Learning Processes: Towards Higher Levels of Automation in Standardized e-Learning**<br>Salvador Sánchez-Alonso and Miguel-Angel Sicilia | 3 |
| **Two Methods for Assessing Critical Thinking in Computer-Mediated Communications (CMC) Transcripts**<br>Patrick J. Fahy | 13 |
| **Learning Statistics in a Shared Virtual Campus: Summarizing a Five-Year Experience**<br>Ana Jesús López and Rigoberto Pérez | 29 |
| **Flying or Falling: Benefits and Pitfalls of Online Peer Review Programs in Distance Education**<br>Kim Blum and Brent Muirhead | 41 |
| **Design and Analysis of Reflection-Supporting Tools in Computer Supported Collaborative Learning**<br>Seung-hee Lee | 49 |
| **Effects of Online Instruction vs. Traditional Instruction on Students' Learning**<br>Yuliang Liu | 57 |

**Editorial**

# The Future of Learning

**Donald G. Perrin**

Education is the most over-regulated and under-funded organization in the modern world. It receives great lip-service and more than its share of criticism. Everyone seems to be an expert on the ills of education because they have been through the process. In order to “fix” the system, politicians, some with very limited educational backgrounds, dictate to educators how to do their jobs. Budgets come with strings attached; parents pressure teachers and administrators; and communities add local political pressures both pro and con.

In a little more than 50 years we have moved from homogeneous K-12 grouping to heterogeneous groups that include a broad spectrum of intelligence levels (IQ), students with disabilities of many kinds - some multiply-handicapped, and students from homes where no English is spoken and communities that are multi-cultural and multi-lingual. We have made every classroom into a one-roomed schoolhouse and expect teachers, both new and old, to meet increasingly diverse student needs in an ever more complex society and simultaneously raise levels of academic achievement.

Even technology is a confounding factor because it is not properly supported. Many teachers lack technology training. Many teacher training institutions have neither personnel nor technology adequate for the task. Big business makes deals with big education for hardware, software, and courseware that may or may not support the needs of teachers and learners. Technology requires training, supplies, maintenance, and periodic replacement – an unlikely scenario in a period of diminishing budgets. And educators are constantly asked to do more - with less.

It would seem we have educational gridlock and as more players join the scrum an appropriate course of action is less clear. Those who are on the front lines – the teachers, administrators, children, and parents, are to be congratulated as survivors. But accolades do not solve problems.

It is time to recognize that the world as we knew it has changed. We are part of a global village. We need new skills to compete and collaborate. We need new models to prepare today’s students for tomorrow’s world. And we need an increasing level of continuing education to be able to grow and change our professions as we go through life.

Distance learning in its many forms is the most malleable technology because it is scalable and transcends many of the traditional barriers of time and distance. It is not surprising that it is being adopted as part of regular academic programs, high schools, elementary schools, home schools, and institutions both public and private.

New tools for customizing learning, such as learning objects, have an obvious role in refining and diversifying curriculum. To be effective, we should add real-world needs assessment, instructional design, flexible delivery, and relevant evaluation - a shift from verbal examination to portfolios, from measurement of knowledge to measurement of skills and performance using criterion based measures or rubricks.

These may not solve the problem of over-regulated and under-funded education, but are intelligent solutions that can make a very significant difference if appropriately applied.

**Editor's Note**: Technology based learning objects continue to be a focus for innovation. This paper adds descriptors to standard Learning Objects (LOs) to facilitate higher levels of automation for e-Learning. It builds on SCORM and similar packaging systems that use Learning Object Management (LOM) to create customized curricula and implement these via a Learning Management System (LMS). This paper shows how addition of descriptors and data to Learning Object metadata can facilitate automation of Learning Object Management and delivery of Learning Objects on Learning Management Systems.

# Normative Specifications of Learning Objects and Learning Processes:

## Towards Higher Levels of Automation in Standardized e-Learning

**Salvador Sánchez-Alonso and Miguel-Angel Sicilia**

## Abstract

Learning object metadata records are nowadays mostly descriptive in the sense that they are intended to give information about the contents or the format of the learning object, but without entailing explicit run-time semantics for Learning Management Systems that use them. Nonetheless, normative metadata descriptions are also required in order to obtain systems that behave according to metadata records in a consistent and predictable way. This paper describes the rationale for normative specification techniques as a complement for existing descriptive metadata, which enables a higher degree of automation by precisely describing usage conditions and expected outcomes for learning objects and learning processes.

**Keywords:** Learning object, metadata, design by contract, ontology, semantic conformance profiles, semantic web, e-learning.

## Introduction

Increasing interest in Web-based education has resulted in standardization efforts to foster portability and shared usage semantics of learning contents and learner information across vendors, platforms and systems (Anido et al., 2002). As a matter of fact, it is possible today to package a Web-oriented course according to standard formats (e.g. according to SCORM[1] packaging models) and then import and use that same content inside any *learning management system* (LMS) that is compliant with the given standard packaging rules. In addition, the scope of such standards and specifications is continuously expanding and covering new areas; for example, the SCORM "sequencing and navigation" specification addresses the standardization of complex navigation and sequencing strategies. Another interesting example is that of IMS "Learning Design"[2], which is targeted to model rich learning activities and their associated pedagogical considerations.

Nevertheless, progress in complexity and coverage of current specifications and standards contrasts with the lack of quality in the level of description of metadata records in existing learning object repositories, which are mostly fragmentary and unstructured, as reported recently

[1] http://www.adlnet.org

[2] http://www.imsproject.org

in a study about the MERLOT repository (Pagés et al., 2003). Najjar's study on the use of metadata in Ariadne (Najjar et al., 2003) also reported that most elements are either never or rarely used by learning object annotators. This study also points out that even those elements, for which values are more regularly provided, are used only in about a fifty percent of the total sum of cases evaluated, except for mandatory items.

In addition to the problem of completeness, current metadata schemas provide room for ambiguity and lack of precision. For example, the LOM standard (IEEE, 2002) – which is consensually considered to be the core of learning object descriptions – may lead to inconsistent usages, since it uses unstructured, natural language fields for many of its elements. For example, LOM category *4.6. "Other Platform Requirements"* is aimed at describing information about software and hardware requirements that can not be expressed by the data element "*4.4.Technical Requirement"*, but it does not make available any value space or guidance about the expected values to be set. In this particular case, the learning object designer has, as the only help available, a pair of vague examples of "other platform requirements" such as "sound card" and "runtime X" (sic). This lack of a clear interpretation is in part due to inexistence of a complete set of consistent vocabularies, what makes most current metadata records unusable for the design and implementation of automated or semi-automated processes like learning object selection, composition or adaptation.

The main problem of LOM from the viewpoint of automation is that it is deliberately *descriptive*, rather than *normative,* with respect to the developer of software processes. Such a descriptive approach is useful for human communication, since human beings are able to understand and even disambiguate descriptions that could hardly be interpreted by current software systems (even though they are equipped with state of the art natural language understanding technology). But descriptive elements do not provide criteria to software systems to drive their actions. In other words, there is not a direct mapping from metadata values to LMS actions that could be used to implement standardized LMS behaviors. For example, how should the "language" metadata element be interpreted? Should it constrain LMS-initiated delivery to students that can proficiently "read" text in the specified languages? A notable exception for this kind of description approaches is the SCORM sequencing specification, which is written in a normative style, since it provides the details of LMS behavior for the user-content interaction. Nonetheless, the core of learning metadata elements is specified in a purely descriptive way.

The "descriptive orientation" cannot be considered as a defect of LOM as a standard, since it does not explicitly target consistent automated behavior as one of its objectives. But it certainly calls for supplementary techniques that fill the gap required to obtain LMSs that act consistently, not only for sequencing, but also for other kind of processes – e.g. composition – which would represent a significant step in standardization of e-learning content and systems.

This paper describes example metadata specification techniques – both for learning object and LMS process descriptions – in a normative style. Such or similar techniques should ideally be integrated with current standards to provide better support for learning management automation, and they would eventually remove the incompleteness and ambiguity of metadata records from annotation practices, by considering metadata completeness and precision as quality metrics for specific usages.

The rest of this article is structured as follows. First examine the current state of metadata standards, focusing on the role of LOM as a metadata communications system between learners and cataloguers. Then, requirements for normative metadata standards and their effects on learning objects are approached. Later, the runtime requirements for LMS processes are described. Finally, the conclusions derived from the previous sections are also provided.

## LOM-Conformant Metadata as a Communications System

The e-learning community is defined in (Wason & Wiley, 2000) as a two-sided scenario where users and cataloguers communicate. While users discover learning objects, probably stored in public repositories, and make use of them in order to attain certain learning objectives, the work of cataloguers consists in tagging educational resources so that users can easily search, find and retrieve resources matching specific criteria. As their communications system is metadata, a consensus needs to have been reached on metadata terms, definitions and values before any fluent communication can start.

Nowadays, LOM has become the most significant and widely used communication system in e-learning applications dealing with Web-based educational resources in the form of reusable learning objects. As a descriptive standard, LOM enables cataloguers to provide metadata values on a number of different aspects, thus allowing users to decide whether a particular learning object is appropriate or not in order to reach a given learning outcome. LOM includes nine different categories covering all the current dimensions of learning objects, but as it is not a closed standard, it can be extended to host future dimensions, as structural, people, relational, etc (LTSC, 2004). Nowadays all the dimensions in LOM are only descriptive, in the sense that LMSs cannot unambiguously adapt or change their runtime behavior depending on the values in the metadata instances.

The lack of a strict formalization in LOM allows cataloguers to set very different values for the same dimension. This situation, and the fact that different cataloguers (or the same one at different moments in time), could provide different values for a given metadata element of a learning object, causes the user-cataloguer communication to be unclear. This is what Wason and Wiley refer to as “noise” in the communications system. Noise is a problem in analog communications that is considerably smaller in digital communications, since here only a set of discrete values can be transmitted. The noise problem in metadata records has been addressed in LOM through the provision of vocabularies that define a set of allowed values for (almost) each metadata element. Unfortunately, vocabularies are not available for all the dimensions in the LOM metadata space, they are not connected to commonsense knowledge representation, and a good number of elements can only be provided values through non-discrete descriptions in natural language.

If LMSs are to behave differently depending on the values of the elements in a metadata record, no uncertainty should ideally be allowed. The definition and use of vocabularies is a promising step towards the definition of precise metadata records, but it does not seem to be enough as to drive LMSs runtime behavior. In fact, ontologies have been recently proposed as substitutes for vocabularies providing richer context descriptions and enabling advanced behaviors – see, for example (Lytras et al., 2003). In addition, it is required that the metadata value establishes its degree of requirement (e.g. mandatory, optional, recommended and the like) and any additional information required like scores or parameters intended to be used by software to act according to them. Normative approaches to metadata and process specification are aimed at covering this latter problem, as described in the following sections.

## Describing normative usage requirements and effects for learning objects

When creating or adapting a given instructional material, learning content designers consider two essential elements as the drivers for the selection of style, interactivity and depth of the contents being developed, namely, the intended audience and the expected learning outcomes (Norman & Nicholson, 1999).

The description of the characteristics of the learner is addressed by "Educational" metadata in LOM, but it does not address fundamental data characteristics that are required if automated matching of learners to learning objects is needed. The following are some of these characteristics:

- The intended LMS usage of some elements.
- The degree of requirement for a description, i.e. whether it is mandatory or optional.
- The degree of credibility for acquiring the expected knowledge or competencies after using the learning object.
- The interpretation of some elements depending on their location in a conceptual representation.

Expected outcomes for a learning object can be of a diverse nature, depending on the effect that the object is intended to drive. Possibly the most common kind of outcome addressed by today's learning objects is knowledge about some kind of subject, involving the development of mental structures. However, the development of abilities is often an objective by itself, and also competencies or social aptitudes (Lave & Wenger, 1991) can be the target of a given learning experience. For example, in a "role play" learning experience simulating a negotiation among different countries, e.g. *The Versailles Experience* described in (IMS, 2003), the negotiation process not only increases each learner's knowledge on the objectives and aims of the rest of the participants, but also provides them with the ability of increasingly improving their negotiation skills. In this respect, even meta-cognitive goals may become the target of a learning object, as described in (Sánchez-Alonso & Sicilia, 2003a).

At the moment, LOM only covers the description of learning outcomes vaguely through elements like *1.5.Keyword*, *1.6.Coverage* and *9.1.Purpose*. However, automation cannot be based only on the learning object expected outcomes as currently defined in LOM. For example, different learners with a different knowledge background could end up attaining different learning objectives after using the same educational resource. Let us consider a learning object on the genitive case in English including examples of advanced use and a final test. Such an object will be more easily assimilated by learners with a sound knowledge of English grammar, even though beginners can also benefit, at least in part, from its use. Current state of metadata specifications doesn't allow learning content designers to clearly state the fact that different users will benefit differently, in terms of learning outcomes, from the use of such an object. Automatic systems or LMSs should consequently not deliver this object to different users on the premise that both will equally benefit from its use, since this is not always true. This situation introduces the need for normative elements that provide learning object designers and authors with the ability of defining standard machine-understandable learning object usage requirements and expected outcomes, which allow automatic or semi-automatic selection of the more appropriate resources depending on the learner's background and other factors.

Normative approaches to learning object metadata should then provide a precise specification of the required behavior of a LMS with regards to each element. Learning object *design by contract* (Sicilia & Sánchez-Alonso, 2003; Sánchez-Alonso & Sicilia, 2003b) is a technique that approaches such normative effect from the viewpoint of contractual relationships between the learning object and the context in which a LMS uses it. This technique basically consists on stating, in the form of declarations called *contracts*, a collection of logical assertions on the requirements of use of a learning object and its expected learning outcomes. Using a recognizable syntax that facilitates automated processing, one or more contracts can be defined for each object, and published for the user community to know about it. Publishing more than one contract for a given object solves the problem of learners without a common knowledge background.

Learning object design by contract redefines the classic correctness formula {P} A {Q} meaning that "any execution of A, starting in a state satisfying P, will terminate in a state satisfying Q" and reformulates it as {C} RLO {O}[θ], to adapt it to the specificities of learning objects. The new meaning is that "the use of the learning object RLO in a learning context C (including a description of specific learner profile) is expected to facilitate the acquisition of the knowledge (or competence or abilities) O [to a certain degree of credibility θ]". In short, a learning object contract looks like:

```
rlo <URI>
      require
            precondition1
            precondition2
            ...
      ensure
            postcondition1
            postcondition2
            ...
```

In this model, preconditions refer to learner profile prerequisites, augmented with platform and other technical and contextual requirements, learning object preconditions stating the constraints under which a learning object can be delivered and used. The syntax of preconditions in contracts, that uses the categories defined by LOM, supports placing information both on the *category* of the requirement: learner, context and system; and on the *level of compromise* of the requirement, which can take the values mandatory, recommended and optional. For example, a precondition stipulating that the system where an object will be delivered must be able to represent text written in Japanese, would be stated in its contract like this:

```
[mandatory] sys.language = jp
```

On the other hand, postconditions are expressed in a syntax that allows learning object authors to include learning outcomes corresponding to different LOM elements. Learning outcomes can be both represented as absolute knowledge attainments, like for example, in the following assertion:

```
lrn.knows(genitiveCaseEnglish)[80]
```

but also as relative to the previous state of the learner's knowledge level, that is represented by '–1'. For example, in a simulation activity aimed at teaching emergency workers on how to handle radioactive waste, learner knowledge will increase every time the learner performs the activity:

```
lrn.knows(handleRadioactiveWaste) > lrn.knows(-1)(handleRadioactiveWaste)[90]
```

In the same way as information about sequencing of learning objects is not part of the metadata, yet introducing attributes that do not describe the content itself, formal information on both the requirements of use and the expected learning outcomes of a learning object could be added to the metadata records as normative attributes.

Other interesting aspect to think about is knowledge conceptualization. Any kind of normative approach to learning object description would ultimately require the presence of some kind of knowledge representations in order to enable richer behaviors than current linear lists of terms (vocabularies) as provided in LOM. Ontologies, understood as conceptualizations that provide an appropriate context for the interpretation of learning object metadata, can be used as:

- A means for the representation of knowledge levels on the learner side.
- A mechanism for the integration of learning object types, essential for the development of systems that are able to select and deliver learning objects. Previous work has addressed this aspect (Sicilia et al., 2004).
- A way to provide reasoning facilities to LMSs, enabled by the underlying description logics (Baader et al., 2003).

This also provides results appropriate for representing postconditions in learning object contracts.

Summing up, a combination of normative descriptions with terminological knowledge representations can be used as the basis for extended learning object metadata specifications to enable a higher level of consistent automation.

## Describing the run-time requirements for LMS processes

The learning processes described so far entail a content-learner (or learner-learner) setting, which can be considered as the "end" process of any LMS. But a high level of automation for learning systems would also expand to other areas that are not constrained to learner participation. In a broad, organizational view of a LMS, it should begin its functioning by some kind of materialization of the "learning needs" of the organization (which is often referred to as "knowledge gap"). Such needs may come from future projects or expected technological changes inside a company, or be part of a formal curriculum. These needs would trigger search processes and selection processes of learning objects. Such selection may involve external providers (ideally, automated learning object repositories) as well as other stakeholders or systems. Several levels of "intelligence" can also be defined to target learning objects and their delivery to the characteristics and time constraints of the employees.

In the broad view of e-learning described here, final delivery and sequencing of learning objects is only a part of the whole process. Standardization should expand its focus to the other "hidden" part of the value chain (Lytras et al., 2002). Much can be borrowed from current B2B specifications like OAGIS[3] or RossetaNet[4], since many learning processes can be considered as business processes.

The notion of "semantic conformance profile" (SCP), described in (Sicilia et al., 2004b), is a recent proposal for definition of learning processes in a broad sense, integrating the ideas of learning object design by contract and pointing to the use of ontological structures as an integral part of definition of processes. For example, the following table summarizes the main elements of a learning object composition profile (CMP-1).

| **Participants** | **Metadata** | | **Run-time Pre-requisites** | **Run-time Commitments** |
|---|---|---|---|---|
| | **Required Elements** | **Idioms** | | |
| The LMS<br>A collection of candidate learning objects $\{LO_i\}$ | LOM (9) *Classifications*<br>Content separation | a) Domain ontology connection with sub-sumption and part of relationships<br>b) Independence | a) Appearance merging.<br>b) Semantic coherence<br>c) Metadata coherence | Matching Algorithm *A*. |

The CMP-1 profile is intended to merge learning objects according to their classification inside taxonomic structures describing their contents. The participants are the LMS making the composition and a collection of candidate learning objects. The presence of LOM Classifications metadata is required, but in addition, such classifications should be connected to an ontological structure which at least represents subsumption (inheritance, "is-a" relationships) and "part-of"

[3] http://www.openapplications.org/

[4] http://www.rosettanet.org/

references. For example, “multi-dimensional arrays in Java” are a part of the subject of “arrays in Java”. These relationships are used in a concrete way described by Algorithm A (which is out of the scope of this paper), so that the behavior of the LMS is an explainable consequence of the annotations regarding Classifications. In addition to that, the learning objects being composed are required to have other properties to be composed together:

- Their contents should be separated from their presentation, so that a given form of “appearance merging” can be done by the LMS. This requirement can be stated in terms of the obligation to use style-sheets.
- The learning objects being composed should be stand-alone (independent), thus not requiring the recursive propagation of the process to other, dependant learning objects. This is done to keep the profile definition simple, and other more complex profile can be defined to define standardized recursive composition in the future.
- Semantic and metadata coherence are required. While semantic coherence can be stated in terms of logical properties of consistence, metadata coherence is more difficult to characterize. For example, the difficulty levels or semantic densities for the learning objects being composed should be compatible (except in the case that differences are explicitly required, but this is not covered in CMP-1).

This way of describing processes internal to LMSs should also be complemented by the definition of a common pattern of messages exchanged between the participants (like they are specified for example, in OAGIS) in case that more than one system is involved, for example, in learning object search, retrieval or publishing processes. In addition, processes are “composable”, in the sense that they can be joined together to form more complex ones. For example, CMP-1 combined with targeting learning objects to specific users (U-SEL) and with a “search into learning object repositories” profile can be considered as a basic profile that fulfills a given learning need inside an organization.

Notations like that of Semantic Conformance Profiles complement normative approaches to metadata with normative description of learning processes of a diverse kind, broadening the scope of current learning technology specifications to the area of system integration.

## Conclusions

Normative approaches to describing learning technology standards and specifications provide the required support to build automated or semi-automated software dealing with diverse aspects of the management of Web-based learning experiences. This is due to the fact that they are oriented to implementers of LMSs that behave in a concrete, predictable way. Since the current basic metadata schemas for learning objects are mostly descriptive, new techniques to complement them in normative styles are required. Learning object design by contract and semantic conformance profiles are two examples of normative techniques based on existing learning technology specifications. The former interprets basic metadata in terms of required conditions and expected outcomes for a learning object. The latter is concerned with a broader view of e-learning. It provides a technique to advance in the normative specification of diverse learning object management processes with a flexible way to specify different levels of complexity and “intelligence” in LMS behavior.

## References


Anido, L. E., Fernández, M. J., Caeiro, M., Santos, J. M., Rodríguez, J. S., Llamas, M. (2002). Educational metadata and brokerage for learning resources. Computers & Education, 38(4), pp. 351 – 374.

Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P. (eds.) (2003). The Description Logic Handbook. Theory, Implementation and Applications. Cambridge.

Carey, T., Swallow, J., Oldfield, W. (2002). Educational Rationale Metadata for Learning Objects. Canadian Journal of Learning and Technology, 28(3) Fall / automne, 2002.

IEEE Learning Technology Standards Committee (2002). Learning Object Metadata (LOM), Final Draft Standard, IEEE 1484.12.1-2002.

IMS Global Learning Consortium, Inc. (IMS) (2003). IMS Learning design best practice and implementation guide.

Najjar, J., Ternier, S., Duval, E. (2003). The Actual Use of Metadata in ARIADNE: an Empirical Analysis. ARIADNE Conference 2003.

Lave, J., Wenger, E. (1991). Situated Learning: Legitimate Peripheral Participation. Cambridge University Press, Cambridge, UK.

LTSC. Joint session with CEN/ISSS Learning Technologies Workshop meeting proceedings, Madrid, 27 January 2004.

Lytras, M., Tsilira, A., Themistocleous, M.G. (2003). Towards the semantic e-Learning: an Ontological Oriented Discussion of the new research agenda in e-Learning. In Proceedings of the Ninth Americas Conference on Information Systems, Tampa, Florida.

Lytras, M., Pouloudi, A., & Poulymenakou, A. (2002). Dynamic e-Learning setting through advanced semantics : The value justification of a knowledge menagement oriented metadata schema. International Journal of e-Learning 1(4), pp. 49-61.

Norman, S., Nicholson, M. (1999). Designing in a Structured World: Instructional Design and Course Development within an SGML/Structured Information Environment - The Open School Experience. In Proceedings of ED-MEDIA 99, World Conference on Educational Multimedia, Hypermedia & Telecommunications, pp. 1448-1449.

Pagés, C., Sicilia, M.A., García, E., Martínez, J.J., Gutiérrez, J.M. (2003). On The Evaluation Of Completeness Of Learning Object Metadata In Open Repositories. In: Proceedings of the Second International Conference on Multimedia and Information & Communication Technologies in Education (m-ICTE 2003), pp. 1760-1764.

Sánchez-Alonso, S., Sicilia, M. A. (2003). Expressing meta-cognitive pre- and post-conditions in learning object contracts. IEEE Learning Technology Newsletter, 5(4), pp. 7-10.

Sánchez-Alonso, S., Sicilia, M. A. (2003). Expressing preconditions in learning object contracts. In Proceedings of the Second International Conference on Multimedia and Information & Communication Technologies in Education, pp. 1656-1660.

Sicilia, M.A., Sánchez-Alonso, S. (2003). On the concept of learning object "Design by Contract". WSEAS Transactions on Systems, 2 (3), pp. 612-617.

Sicilia, M.A., García, E., Sánchez-Alonso, S., Rodríguez, E. (2004). Describing learning object types in ontological structures: towards specialized pedagogical selection. In Proceedings of ED-MEDIA 2004 - World conference on educational multimedia, hypermedia and telecommunications.

Sicilia, M.A., Pagés, C., García, E., Sánchez-Alonso, S. (2004). Specifying semantic conformance profiles in reusable learning object metadata. In Proceedings of the 5th International Conference on Information Technology Based Higher Education and Training: ITHET 2004.

Wason, T., Wiley, D. (2000). Structured Metadata Spaces. Journal of Internet Cataloging, volume 3 (2/3), pp. 263-277.

## About the Authors

**Salvador Sanchez-Alonso** obtained a university degree in Computer Science from the Pontifical University of Salamanca (Spain) in 1997. He worked as an assistant professor at the Pontifical University of Salamanca in Madrid from 1997 to 2000 and from 2002 to 2005. He also worked as a software engineer at a software solutions company during 2000 and 2001. From 2005, he is a professor of the Computer Science Department of the University of Alcala, in Spain. Currently finishing his PhD thesis on the design and use of contract-based learning objects, his research interests include Learning objects reusability, Metadata, Object-Oriented Technologies and Web Engineering.

E-mail contact: salvador.sanchez@uah.es

**Miguel A. Sicilia** obtained a university degree in Computer Science from the Pontifical University of Salamanca, Madrid, Spain, in 1996 and a Ph.D. degree from the Carlos III University in 1999. From 1997 to 1999 he worked as an assistant professor and later on as a part-time lecturer at the Computer Science Department of the same university. He also worked as a software architect in e-commerce consulting firms. From 2002 to 2003 he worked as a full-time lecturer at the Carlos III University, after which he joined the University of Alcalá. His research interests are primarily in the areas of adaptive hypermedia, learning technology, and human-computer interaction, with a special focus on the role of uncertainty and imprecision handling techniques in those fields.

E-mail contact: msicilia@uah.es

Both authors can be contacted at:

University of Alcalá, Computer Science Department, Ctra.
Barcelona km 33.6 – 28871 Alcalá de Henares (Madrid), Spain

**Editor's Note**: This research combines two instruments to detect aspects of critical thinking in online interactions. It explores theoretical and practical constructs for mentoring discussions, analysis of critical thinking processes, and interpreting the findings. By codifying each sentence in an online discussion, it is possible to generate statistics and descriptors for the critical thinking process – trigger, exploration, integration, and resolution. These findings have value for design and implementation of online learning and mentoring.

# Two Methods for Assessing Critical Thinking in Computer-Mediated Communications (CMC) Transcripts

**Patrick J. Fahy**

## Abstract

Critical thinking, though critical in education, is especially difficult to detect in online learning and teaching based on computer-mediated communication (CMC). As a latent construct, critical thinking must be inferred by analysis of the "traces" of higher-level cognitive activity found in transcripts. Two models are presented for describing and analyzing critical thinking, the *practical inquiry* (PI) model (Garrison, Anderson, & Archer, 2001), and the *Transcript Analysis Tool* (TAT) (Fahy, Crawford, & Ally, 2001). The models reveal different aspects of online interaction: the PI model determines the proportions of four phases found in transcripts of the critical thinking process, while the TAT adds detail, from the sentence level, about communication strategies and patterns within postings. Principal findings and suggestions for further research focus on *triggers* and postings classified as *other* in the PI model.

**Keywords**: CMC; online interaction analysis; transcript analysis; community of inquiry; cognitive presence; practical inquiry model; Transcript Analysis Tool (TAT); network interaction; social presence

## Background

Critical or higher-order thinking has consistently been cited as a prime objective of all types of education, including education at a distance (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956; Gibson, 1996; Bostock, 1997; Romiszowski, 1997; Haughey & Anderson, 1998; Marttunen, 1998; Collison, Elbaum, Haavind, & Tinker, 2000; Strother, 2002; Roblyer & Schwier, 2003). In pursuit of a better understanding of this critical construct, the Canadian research group of Garrison, Anderson, Rourke, and Archer have articulated, in three important papers, a conceptual framework for the context within which they argue critical thinking is likely to be found, a *community of inquiry* (Rourke, Anderson, Garrison, & Archer, 1999; Anderson, Rourke, Garrison, & Archer, 2001; Garrison, Anderson, & Archer, 2001). In their paper of interest here (Garrison, et al., 2001), they stated that such a community, engaging in critical thinking, is an "extremely valuable, if not essential" element of higher education. Their work focuses on the importance, to communities of inquiry, of computer-mediated communications (CMC) as a means for creating and sustaining cognitive presence, and as a vehicle for engaging in critical thinking (Gunawardena, Lowe, & Anderson, 1997; Fahy, 2001).

The efforts of Garrison et al. (2001) are important, as they directly address a fundamental problem encountered by all attempts to detect or assess latent constructs such as cognitive presence and critical thinking in online contexts (Rourke & Anderson, 2004). Because they are only indirectly knowable, latent constructs must be known from their "traces," "symptoms," or "indicators" (2001, p. 12, 13). This process, as the authors admit, is inherently subjective, inductive, and prone to error (p. 12) – as is, one might add, the associated research.

This paper was prompted by two objectives. The first was a wish to apply further, and perhaps refine, Garrison et al.'s (2001) *practical inquiry* (PI) *model* (the term for the operationalized procedure for using this construct, grounded in the critical thinking literature; p. 8). Garrison et al. conducted an admittedly weak initial pilot application of their model, commenting that it should "not be seen as immutable" (p. 9), and concluding their paper with the comment, "this tool is worth further investigation" (p. 22). A review of the literature since its appearance reveals that to date the model has not received the further testing the authors hoped and expected it would, a situation, it is hoped, this paper will in part redress.

A second reason for this paper was to present a comparison of Garrison et al.'s method for the detection and assessment of latent projective variables (including, but not limited to, cognitive presence; Fahy, 2003) with that of another model, the *Transcript Analysis Tool* (TAT) (Fahy, Crawford, & Ally, 2001). The comparison is guided by Thorngate's (1976) *postulate of commensurate complexity* (cited in Weick, 1979, p. 35ff.): "It is impossible for a theory of social behavior to be simultaneously general, accurate, and simple" (p. 35). The TAT model requires each sentence be considered (coded). In comparison to the PI model, the TAT strives for accuracy and generalizability, at the expense of the reliability that greater simplicity would confer. In presenting the PI model, Garrison et al. wrote of their intention to avoid some of the complexities of other approaches, for example by restricting the number of phases (and therefore the number of coding distinctions) to four, a relatively low number in transcript analysis studies (Fahy, 2001; Rourke, Anderson, Garrison, & Archer, 2001); by coding whole postings rather than component parts; and by developing a "heuristic" for dealing with "contradictory categorization cues or evidence of multiple phases" (p. 17), that required "coding up" or "coding down" (to resolve situations where coding was ambiguous). While their approach produced good reliability figures, the authors recognized the implications of their approach in relation to Thorngate's third criterion, *accuracy*: "Submessage level units [i.e., sentences] may be introduced in future confirmatory studies if increased precision is warranted" (p. 17).

The importance of critical thinking as a component and outcome of online interaction has piqued previous research interest. Various attempts have been made to operationalize critical thinking in order to see it more clearly in online group behaviour, occasionally (as will be seen in the following list) simply by attributing it to certain activities or to specific strategies or technologies (Simon & Berstein, 1985). Examples of de facto perception of critical thinking to activities and tools include: questioning and challenging (Blanchette, 2001); constructivist dialogue in case-based learning contexts (Commonwealth of Learning, 1993; Jonassen, Davidson, Collins, Campbell, & Bannan Haag, 1995; Jonassen, 1998); collaborations of various forms and under different conditions (Bullen, 1998; Curtis & Lawson, 2001; Rose, 2004); group focus and reflection on transcript contents (Davie & Wells, 1992); uses of various media (Dede, 1996; Mayer, 2001); and approaches to group-mediated strategic thinking (Gunawardena, et al., 1997).

The above shows how widely researchers have ranged to find evidence of cognition in CMC. In this paper, the two methods used to detect critical thinking in an online community focus on "the nature and quality of critical discourse" found in the transcript itself. While their methods differ, the purpose of both methods is to identify elements of postings that create identifiable and (more or less) predictable responses from others in the online community. Both models address the *organization* of online interaction, "a systematic account of some rules and conventions by which sets of interlocked behaviours are assembled to form social processes" (Weick, 1979, p. 3). They focus on different elements of the transcript (postings in the Garrison et al. model, sentences in the TAT), but the intention in both cases is a "systematic account" of some aspect of communication in the online community, derived directly from transcript data. Their use together here is exploratory; consequently, parts of the analysis assume that exploratory studies may claim some latitude in interpretation in the interests of fairly testing their potential value (Rourke, et al., 1999).

## Two models for detecting critical thinking in online interaction

*The Practical Inquiry model.* Garrison et al. (2001) operationalized critical thinking through a *model of practical inquiry*, recognizing such thinking to be both a process and an outcome of online communities engaged in reflective critical discourse (p. 7, 8). In *critical communities of inquiry*, they reasoned, participants apply *reflection* and *action* to facts and ideas, often (especially in educational environments) under the direction of a moderator or instructor. (This concept is similar to that of McKlin, Harmon, Evans, & Jones [2002, p. 2], who linked "sustained reflection and discourse" to cognitive activity.)

The phases of this model of critical thinking are as follows (Garrison et al., 2001, pp. 10 – 11):

- A **triggering event** begins the inquiry process. A trigger is a problem or dilemma, usually initially defined or identified in educational situations by the instructor/moderator. The process includes identifying and focusing on one trigger (sometimes explicitly rejecting or excluding others).
- **Exploration** involves movement between the private, reflective world, and the shared, collaborative world, with participants alternating from reflection to discourse as they strive to grasp or perceive the problem and understand its nature. This phase is typified by brainstorming, questioning, and free exchanges of information. The authors warn, that students may resist moving out of this phase into the next unless prodded by the instructor/moderator.
- **Integration** is the phase where meaning is constructed from the ideas generated in the previous phase. Ideas are evaluated on the basis of how well they connect with and describe the problem. Participants may continue to move repeatedly from private reflection to public discourse in this phase of the inquiry process. This is the most difficult phase to detect – its presence must often be inferred from other evidence.
- **Resolution** is signified by the appearance of vicarious or direct action. In non-educational situations, this is often in the form of actual application of the solution; in educational contexts, tests or applications are usually vicarious or hypothetical. Resolution requires "clear expectations and opportunities to apply newly created knowledge" (p. 11). If the resolution is perceived as incomplete or inadequate in any way, or a new problem is identified, the process may be repeated.

The PI model was initially tested by its developers on a corpus of 95 postings, small enough to be called by the authors a "methodological weakness" (Garrison, et al., 2001, p. 18). The researchers coded messages from the transcript using the four original categories (soon adding a fifth category, *other*, to accommodate messages not fitting elsewhere). When (perhaps not unexpectedly; see Fahy, 2001; Rourke, et al., 2001; Puustjärvi, 2004) difficulties were encountered with classification of whole postings into single categories, further clarification in the form of "descriptors" and the "perspective" of the participants was added to assist coding. The phases (codes), with the adjunctive "descriptors" and "perspectives," are shown below.

**Table 1**

**Phases, descriptors, and perspectives of the Practical Inquiry model**

| Phase | Descriptor | Perspective |
|---|---|---|
| *Triggers* | Evocative | Shared world |
| *Exploration* | Inquisitive | Private world |
| *Integration* | Tentative | Reflection |
| *Resolution* | Committed | Discourse |
| *Other* | (Postings not fitting another category) | |

The validity of the process appeared promising in the initial application: codings in the pilot test of the model yielded coefficients of reliability (*CR*, a ratio of agreement to total number of judgments made by raters) ranging from 0.45 to 0.84, and *kappa* values from 0.35 to 0.74 (Garrison, et al., 2001, p. 18). (Cohen's *kappa* is a chance corrected measure of agreement [University of Colorado, 1999; Agreement observer, 2000], especially useful where the number of coding decisions is limited, thus making chance a potentially important factor in the classification process).

The results of the initial pilot analysis (see Table 2, in "Findings," below) showed that most of the postings (42%) were exploration, and that the next most common category (consisting of eight postings, or one-third of the total) was *other,* postings that could not be classified in any of the other four phases (p. 19). In the pilot test of the model, the authors wrote that their intention in offering the PI model was to suggest an approach that might be useful in facilitating the process of higher-order online learning (Garrison et al., 2001, p. 8), and that the model's phases reflected an "idealized" critical inquiry process which "must not be seen as immutable" (p. 9), words encouraging to the present study.

<u>*The TAT model.*</u> Another approach to understanding the content and social processes in online interaction, including thinking processes, is the TAT (*Transcript Analysis Tool*). The TAT, based on a concept originated by Zhu (1996), has been applied during its development to a variety of CMC-based interaction analysis problems (Fahy, et al., 2001; Fahy, 2002a; Fahy, 2002b; Fahy, 2003; Fahy, 2004; Fahy & Ally, in press). Application of the TAT involves coding each sentence of a transcript into one of 8 categories (five major): 1) *questions* (horizontal or vertical), 2) *statements* (referential or non-referential), 3) *reflections,* 4) *scaffolding comments,* or 5) *paraphrases and citations.*

Briefly, the categories and designations of the TAT are as follows:

## Type 1 - Questions:

**1A** includes *vertical questions*, which assume a "correct" answer exists, and the question can be answered if the right (knowledgeable) individual is asked, or the right source contacted.

**1B** are *horizontal questions*: recognizes there may not be one right answer; others are invited to help provide a plausible or alternate "answer" or explanation, or to help shed light on the question.

## Type 2 - Statements:

**2A (non-referential statements)** contain little self-revelation and usually do not invite response or dialogue; the main intent is to impart facts or information. The speaker may take a matter-of-fact, didactic, or pedantic stance, providing information or correction to an audience assumed to be uninformed or in error, but curious, interested, and otherwise open to correction. *Statements* may contain implicit values or beliefs, but usually these must be inferred, and are not as explicit as they are in *reflections* (TAT type 3).

**2B (referential statements)** comprise direct answers to questions**,** or comments making reference to specific preceding statements**.**

### *Type 3 - Reflections* (significant personal revelations)

***Type 3*** sentences show the speaker expressing thoughts, judgments, opinions, or information which are personal and are usually guarded or private. The speaker may also reveal personal values, beliefs, doubts, convictions, or ideas acknowledged as personal. The listener/reader receives both information about some aspect of the world (in the form of opinions), and insights into the speaker. Listeners are assumed to be interested in and empathetic toward these personal revelations, and are expected to respond with understanding and acceptance. The speaker implicitly welcomes questions (even personal ones), as well as self-revelations in turn, and other supportive responses.

### Type 4 - Scaffolding/engaging

Scaffolding/engaging sentences are intended to initiate, continue, or acknowledge interpersonal interaction, to “warm” and personalize the discussion by greeting or welcoming, and to support and maintain the online network by enhancing inclusiveness. Scaffolding/engaging comments connect or agree with, thank, or recognize someone else, and encourage or acknowledge the helpfulness, ideas and comments, capabilities, and experience of others. Also included are comments without real substantive meaning (“phatic communion,” “elevator/weather talk,” salutations/greetings, and closings/signatures), and devices such as obvious rhetorical questions and emoticons, whose main purpose is maintenance of the interpersonal health of the online community.

### Type 5 - Quotations/citations:

**5A**: quotations or paraphrases of others’ words or ideas, including print and non-print sources.

**5B**: citations or attributions of quotations or paraphrases, in a formal or reasonable complete informal manner.

The TAT uses sentences; each sentence in the transcript is assigned to one (or more) TAT categories (about 6% of sentences in this transcript received more than one TAT code, a typical proportion). *Unitizing*, the process of selecting elements of the transcript to code, has sometimes proven problematic (Rourke, et al., 2001; Fahy, 2001). While the debate has not been resolved, problems have been identified with units greater than the sentence, such as “units of meaning” (Henri, 1992), “segments” (Borg & Gall, 1989, cited in Garrison et al., 2001), “thematic units” (Rourke, et al., 1999), or “phases” (Gunawardena, et al., 1997). Although Garrison et al. coded their transcript at the level of the posting, for reasons of consistency and due to concern for validity, they acknowledged (2001, p. 17), as noted earlier, the advantages of sentence-level analysis for revealing more accurately subtle nuances in the transcript (Fahy, 2001, 2002a, 2002b).

## Theoretical context for an analysis of critical thinking in CMC

Garrison and his colleagues posited in the PI model that critical thinking would involve a progression through four phases, beginning with a *trigger*, moving through *exploration*, to *integration*, and achieving final *resolution*. They reasoned that higher-order learning required questioning and challenging of assumptions, through the dual processes of engagement in internal *reflection* and community-based *discourse* (via CMC), resulting in further (re)constructing of experience and knowledge. Critical thinking, in this view, requires interaction with a community, drawing upon the resources of the community to test the content of individual contributions (the quality of ideas, the soundness of reasoning, the universality of experience, cogency of argument, eloquence, etc.).

In proposing four main phases for this process, the PI model presents a cyclical concept of thinking (resolution, the final phase, may reveal new dichotomies or discontinuities, producing a new triggering event); in general in this model, groups are assumed to be seeking resolution. While each phase of the model is accompanied by concurrent cognitive and social outcomes, the implication is clear that the overall process is incomplete if it stalls prior to completion of a full cycle ending with resolution (p. 9).

The initial pilot application of the PI model revealed little integration (Table 2), and even less resolution (Garrison, et al., 2001, p. 18). This finding may not be surprising, for theoretical reasons which others (including one of Garrison's co-authors) have identified. Kanuka and Anderson (1998) examined a transcript generated in a moderated online forum (CMC conference), whose purpose was to support professional development among distance education professions. The researchers sought evidence of a five-phase knowledge-construction process, based on constructivist theory:

1. Sharing and comparing information;
2. Discovery and exploration of dissonance or inconsistency;
3. Negotiation of meaning/co-construction of knowledge;
4. Testing and modification of proposed synthesis or co-construction;
5. Phrasing of agreement, statement(s), and applications of newly constructed meaning.

In fact, about 93% of the transcript postings (191 of 216; an "overwhelming number," according to the authors [p. 65]) fell into the first category. This single phase, *sharing/comparing of information*, as defined in the study by the researchers, consisted of various preparatory activities, including several reminiscent of those found in the *triggering* phase of the PI model: stating observations or opinions, expressing agreement or support, identifying problems, defining, describing, corroborating, and clarifying questions. The other four phases, including especially those equivalent to what Garrison et al. termed *integration* and *resolution*, comprised as little as 3% of the transcript, depending upon the proportion deemed *exploration* (Kanuka & Anderson, 1998, p. 66). These results suggested that the analytic approach used in the study may not have discriminated adequately to permit real insights into the quality of the online interaction, a previously described problem in transcript analysis studies (Fahy, 2001; Rourke, et al., 2001).

Another example was reported by Gunawardena, et al. (1997). Using a similar analytic approach, they attempted to use the structure of a stringently moderated online debate to examine the social construction of knowledge in an international group of experienced distance education professionals. The authors held that knowledge results from interaction, stating emphatically: "Interaction is the process through which negotiation of meaning and co-creation of knowledge occurs" (p. 405). They assumed knowledge construction would occur in this group despite the debate structure, since the interaction was collaborative as opposed to one-way (p. 400 - 401).

Of particular interest in this study was the finding that participants obviously resisted the debate format, attempting to reach compromise and consensus despite the persistent efforts of the debate leaders "to keep the two sides apart" (p. 417). In effect, the researchers reported, the moderators' attempts to base the discussion on discord ran counter to the group's preference for synthesis.

Even in a formal debate, these findings showed, the group's propensity may be to avoid dwelling on differences, and to seek commonalities.

The work of Fulford and Zhang (1993) may partially explain these findings. Fulford and Zhang studied perceptions of interaction among teachers involved in professional development, by examining the interaction of the variables personal interaction, overall interaction, and satisfaction. The findings of interest were, first, that perceptions of personal and overall interaction were positively correlated ("people who see themselves as active participants tend to have a more positive perception of overall interaction" [p. 14]); second, that satisfaction was more attributable to perceived *overall* interactivity than to *individual* participation, leading to the conclusion that "learners who perceive interaction to be high will have more satisfaction with the instruction than will learners who perceive interaction to be low" (p. 18). An encouraging and intriguing finding for instructor/moderators was the observation that involving all students in direct instructor-student interaction might not be necessary to produce positive perceptions of overall group interactivity: "Vicarious interaction may result in greater learner satisfaction than would the divided attention necessary to ensure the overt engagement of each participant [by the instructor]" (p. 19).

The above suggests that in their cognitive behaviours online groups may have a disposition (a tropism, in biological terms) toward consensus, agreement, synthesis, and accord, and an aversion to discord, conflict, and argument. Rather than seeking a clash of viewpoints in CMC, participants apparently prefer to attempt to build solidarity. As Gunawardena, et al. noted, in group interactions "the situation itself exerts a strong mediation effect upon individual cognitive and conceptual processes" (p. 407), favouring sharing and concord. The relative lack of conflict in instructor-moderated academic interactions, especially in comparison with the Mardi Gras-like atmosphere often seen in unmoderated list-based discussions (Walther, 1996; Yates, 1997; Schrage, 2003), may be seen as further evidence of this preference (Garton, Haythornthwaite, & Wellman, 1997).

The finding of Garrison et al., (2001), Kanuka and Anderson (1998), and Gunawardena, et al. (1999), that online groups appear "comfortable remaining in a continuous exploration mode" (Garrison, et al., 2001, p. 10), requiring moderator intervention (or "teaching presence"; Anderson, et al., 2001) to move to more advanced stages of critical thinking, is one of several generalities following from these studies. Others include:

> For individuals, the process of critical thinking involves both private reflection and public interaction, the latter within a community;
>
> Efforts to observe interaction associated with critical thinking often produce results which do not discriminate well (a few interaction categories [codes] account for a large proportion of the observations), or expose weak or faulty instruments, or poor observational procedures;
>
> CMC participants engaged in a process of critical thinking seem to prefer to share and compare, and to avoid conflicts, differences of opinion, or disagreements of interpretation;
>
> The tendency to avoid overt disagreement and discord may be based on a group preference for a climate where the quality of general social interaction is more important to satisfaction than opportunities for personal interaction (a climate that is more *epistolary* than *expository*) (Fahy, 2002a).

This present study was designed to explore the behaviour of an online community engaged in critical thinking, as reflected in the transcript of its online CMC interactions, by the application of two different but similarly purposed analytic models. The portion of the *total* intra-group interaction that occurred is not known, as students had the option of communicating by other means not assessed in the study (e-mail, telephone, even face-to-face meetings). The assumption here, as in similar studies, was that the transcript would contain evidence – "traces" (Garrison et al., 2001, p. 12) – showing how the community of inquiry was functioning as a unit in relation to its sociocognitive purposes, and that these two tools would reveal important, but different, elements of that functioning.

## Method

The study corpus used was a transcript of 462 postings, comprising 3,126 sentences containing approximately 54,000 words, generated by a group of thirteen students and an instructor/moderator, engaged in a 13-week distance education graduate credit course delivered totally at a distance. All of the students were experienced CMC users, and the instructor was an experienced distance educator who had used CMC to instruct graduate courses at a distance for over five years.

Each posting of the study transcript was coded into one of the PI model's categories (*trigger, exploration, integration, resolution*); each sentence was also coded with the TAT (5.3% of the sentences received more than one TAT code). A code-recode method was used: the author did the initial coding of the transcript using both models, then recoded it again more than two months later. For the TAT, coefficient of reliability (CR) values ranging from of .70 to .94 have been reported (Keller, 1999; Fahy, Crawford, Ally, Cookson, Keller, & Prosser, 2000; Fahy, Crawford, & Ally, 2001; Poscente, 2003). In this case, the agreement level (*CR*) was 81% with the TAT (Fahy, et al., 2001).

For the PI model, the whole posting was coded into one of the model's five categories. As noted above, the process of fitting whole postings into one code can be problematic: postings often contain multiple elements, and forcing a whole post into one category may ignore nuances or shadings of meaning. The PI model's authors recognized this problem, recommending "coding down" to an earlier phase when it is not clear which phase is reflected, and "coding up" to a later phase when evidence of multiple phases was detected (Garrison et al., 2001, p. 17). (The frequency with which coding up or down was applied was not reported in the original paper.) In this study, coding up and down was applied as described when required, and an overall code-recode reliability of 86% was achieved with the PI model.

Coding for both models was accomplished with ATLAS.ti, and quantitative analyses were conducted with SPSS-PC and Excel.

## Findings

Table 2 shows the results obtained from the application of the PI model to the study transcript, compared to the findings reported from the initial small pilot implementation of the model at the time of its initial appearance (Garrison, et al., 2001).

As shown in Table 2, while the proportions of postings in the categories of *trigger, integration, and integration/resolution* are remarkably similar in both studies, *exploration* was clearly affected by the large difference in the postings coded as *other*. In the original study, the process of coding three transcripts to refine the process produced interrater reliabilities from .45 to .84 (Garrison, et al., 2001, p. 18); the most frequent interrater disagreement during the refinement process reportedly occurred between the phases *exploration* and *integration* (p. 19). As well, during development and refinement of the model the category of *other* was added to the initial four phases; by the third transcript coding there was no reported disagreement among the coders in identifying postings placed in this category (p. 19).

**Table 2**

**Practical Inquiry (PI) model results**

| Phases of the PI model | Initial pilot | | Present study | |
|---|---|---|---|---|
| *Phase* | # | % | # | % |
| Trigger | 2 | 8 | 42 | 9.1 |
| Exploration | 10 | 42 | 331 | 71.6 |
| Integration | 3 | 13 | 65 | 14.1 |
| Resolution | 1 | 4 | 8 | 1.7 |
| Other | 8 | 33 | 16 | 3.5 |
| **Total** | 24 | 100 | 462 | 100 |

Table 3 shows the occurrence of TAT categories, at the level of the sentence, within each of the five phases of the PI model.

**Table 3**

**TAT Results**

| TAT sentence type | Trigger | Explore | Integrate | Resolve | Other | Total | S.D. |
|---|---|---|---|---|---|---|---|
| 1A – Horizontal question | 1.0% | 1.9% | 2.0% | 1.3% | 2.2% | 1.9% | 0.49% |
| 1B – Vertical question | 21.3 | 1.8 | 1.5 | 2.6 | 0.0 | 3.0 | 8.92 |
| 2A – Non-ref. statement | 34.5 | 52.2 | 41.4 | 46.1 | 45.7 | 49.3 | 6.53 |
| 2B – Referential statement | 6.1 | 10.0 | 12.6 | 11.8 | 13.0 | 10.2 | 2.83 |
| 3 – Reflection | 8.1 | 19.3 | 18.3 | 22.4 | 2.2 | 18.3 | 8.53 |
| 4 – Scaffolding statement | 10.7 | 9.6 | 16.7 | 11.8 | 37.0 | 11.2 | 11.40 |
| 5A – Quotation, paraphrase | 10.7 | 3.5 | 5.1 | 1.3 | 0.0 | 4.1 | 4.15 |
| 5B - Citation | 7.6 | 1.7 | 2.4 | 2.6 | 0.0 | 2.1 | 2.85 |
| Number of sentences | 197 | 2353 | 454 | 76 | 46 | 3126 | |
| Total (%) | 6.3 | 75.3 | 14.5 | 2.4 | 1.5 | | |

A comparison of Tables 2 and 3 shows some small discrepancies in the proportion of sentences (Table 3), compared with the frequency of the phases (Table 2): while *triggers* constituted over 9% of the phases, they comprised only 6.3% of the sentences; *exploration* tended to contain more sentences than its proportion of the phases (75.3% vs. 71.6%, respectively); *integration* was almost identically in proportion (14.1 of phases and 14.5% of sentences); *resolution* contained a higher proportion of sentences than its share of phases (2.4% vs. 1.7%); and *other* postings, while comprising 3.5% of the phases, constituted only 1.5% of the sentences. The pattern suggests that *triggers*, *resolution*, and *other* postings tended to be shorter (in numbers of sentences), while *exploration* and *resolution* postings tended to be lengthier. This finding is not surprising: one would expect that the processes of exploring and achieving resolution of issues would require more interaction (as seen in the number of sentences), while initiating the process, or comments orthogonal to the topic, would require less.

In order to provide a standardized method of assessing the proportions observed in Table 3, and to identify potentially salient findings for further investigation in this exploratory study, *z* (standard) scores were calculated. The *z* statistic shows the distance of the figure of interest (in this case, the percentages shown in Table 3, reflecting the proportion of TAT sentences within each phase) from the mean, in standard deviation units (Best, 1970). Table 4 shows the *z* scores for these percentages. (Cells of interest in relation to the following discussion are shown left-aligned and in **bold** in the following Table.)

**Table 4**

**TAT results converted to *Z* scores**

| TAT sentence type | Trigger | Explore | Integrate | Resolve | Other |
|---|---|---|---|---|---|
| **1A – Vertical question** | -1.71 | 0.12 | 0.26 | -1.10 | 0.65 |
| **1B – Horizontal question** | 2.06 | -0.13 | -0.16 | -0.04 | -0.33 |
| **2A – Non-referential statement** | -2.26 | 0.45 | -1.20 | -0.49 | -0.55 |
| **2B – Referential statement** | -1.47 | -0.07 | 0.82 | 0.57 | 0.99 |
| **3 – Reflection** | -1.19 | 0.12 | 0.00 | 0.48 | -1.89 |
| **4 – Scaffolding statement** | -0.04 | -0.14 | 0.49 | 0.06 | 2.26 |
| **5A – Quotation, paraphrase** | 1.59 | -0.14 | 0.24 | -0.66 | -0.98 |
| **5B – Citation** | 1.92 | -0.17 | 0.10 | 0.17 | -0.75 |

As can be seen, the phase with the greatest TAT variations was *trigger* postings, while the least variation was found in *exploration* postings. As described below, the phase *other* also contains some intriguing findings. The following summarizes the differences noted in the Table. (For this exploratory study, a *z* score of ±1.5 standard deviations is termed *salient*, while a difference of ±2.0 S.D. is considered *significant*).

Table5 summarizes the findings in relation to the TAT analysis, for significant and salient results.

**Table 5**

**Summary of differences in TAT sentence types within PI phases**
**($z \geq 1.50$)**

| PI Phase | Valence | TAT Category | Effect Size ($z$ score) |
|---|---|---|---|
| **Triggers** | More: | Horizontal questions (1B) | 2.06 |
| | | Citations (5B) | 1.92 |
| | | Quotations and paraphrases (5A) | 1.59 |
| | Fewer: | Non-referential statements (2A) | -2.26 |
| | | Vertical questions (1A) | -1.71 |
| **Other** | More: | Scaffolding/engaging (4) | 2.26 |
| | Fewer: | Reflections (3) | -1.89 |

Most *triggers* originated with the instructor/moderator, in accord with the predictions of Garrison et al. (2001): in the study transcript, 74% of the *trigger* postings were made by the instructor/moderator, 26% by students. This was the only phase where such a marked difference was noted, and conforms to the description of *triggers* in the PI model as a primary pedagogical responsibility of the instructor/moderator.

Four other findings in Table 3 are discussed here briefly, as suggestive in relation to the significant and salient findings reported earlier (the *z* scores associated with these differences were less than 1.5, but were in the same direction as the other findings, perhaps warranting further investigation (Riffe, Lacy, & Fico, 1998, in Rourke, et al., 1999, p. 66). In relation to *triggers,* two other TAT categories were also less common: *referential statements* ($z$ = -1.47) and *reflections* ($z$ = -1.19). Added to the previous significant and salient findings, these suggest *triggers* may also comprise more *horizontal questions*, *quotations/paraphrases,* and *citations*, and less of the other TAT categories, a finding similar to Poscente's (2003).

*Integration* was also found to contain a somewhat lower proportion of *non-referential statements* (2A; $z$ = -1.20), with a slightly elevated level of *referential statements* (2B; $z$ = 0.82). These differences support a view of *integration* as a phase of interactive construction of meaning, involving assessing, connecting, and describing emerging understandings (Garrison, et al., 2001, p. 10), through both referential and non-referential statements.

Finally, *resolution* contained fewer *vertical questions* (1A; $z$ = -1.10). As the phase in which consensus is built by vicarious or actual application of the knowledge developed in the other phases, this fact, and the presence of somewhat more *referential statements* and *reflections* (Table 4), are together not unexpected.

The above analysis permits the following summary of the nature of the online interaction observed here:

The frequencies of the PI model's phases were similar to those noted in the original report, with the bulk of all postings constituting *exploration*, and *triggers* and *integration/resolution* comprising much smaller proportions of the interaction.

The contents of the category *other* in the PI model warrants further investigation, especially in regard to the apparently greater *social* and *network* orientation of this phase (revealed by the slightly higher proportion of *scaffolding/engaging* sentences).

The TAT analysis showed a tendency in *exploration* and *resolution* postings for more sentences, and in *triggers* for fewer. (The relation of posting length to type or contents remains unresolved, and in need of further study.)

On the basis of relative differences among TAT categories, revealed by *z* scores, *triggers* differed most from the other phases in terms of the TAT constituents, containing significantly more *horizontal questions, quotations and paraphrases,* and *citations*, and significantly fewer *vertical questions* and *non-referential statements*.

# Discussion

The two different approaches to the analysis of the same study transcript revealed different aspects of the kind and quality of the online interaction that generated it. The PI model showed similar relative proportions of most of the phases as were found in an initial application, but the reduced occurrence of the phase *other* raises questions about the nature of this category, and about activities within the online community itself. The task of analysis was made more difficult by the fact that little information was provided regarding the type of postings which were classified *other* in the original work; the discrepancy found here could therefore be due to a lack of agreement about what *other* comprises (resulting in this study in the coding into one of the four principal phases material that was not coded that way by the authors of the original study), or it may reflect a genuine difference between this transcript and the one used by Garrison et al. (2001) in their initial paper.

*Other* comments are inherently difficult to classify, being defined by what they are *not* (one of the other four phases). A clue to the nature of these postings, and to a fundamental difference in the two analytic approaches, was the significantly higher occurrence in *other* postings of TAT *scaffolding/engaging* sentences, the type which addresses network maintenance and inclusiveness in the online community. These may indicate that the PI model does not provide for such factors within its four main phases. The fact that the TAT was able to identify the greater presence of the *scaffolding/engaging* sentence type suggests a difference, and perhaps an advantage, in relation to detection of specific kinds of interpersonal content in transcripts. These results are preliminary; further studies are clearly needed, carefully examining coding decisions relating to the *other* category. (Garrison et al. commented, "Content analysis is a difficult process under the best of circumstances" [2001, p. 18]; one suspects that grappling with complexities such as *other* content might have prompted that observation.)

Other findings at the level of the sentence seemed to confirm that the TAT and the PI model were both sensitive to similar processes within postings, and that these processes were consistent with their notional designations. This was especially evident in regard to *triggers.* In the PI model, *triggers* are *sui generis*, initiated by the instructor/moderator to focus group attention on a problem or phenomenon. In this study, the task of triggering the group was clearly one predominantly – though not exclusively – exercised by the teacher/moderator, and this pattern was detected equally well, although in different ways, by both tools.

Characteristics of *integration* and *resolution* postings were also revealed by the dual analysis. First, there was some evidence of reliability: the proportions of these two phases were found to be similar in both studies. Second, somewhat lower levels of *non-referential statements* and *vertical* questions were found in these phases, accompanied by more *referential statements*. These interactive processes, made apparent by the TAT analysis, may be the actual communicative strategies, or linguistic "moves" (Herring, 1996), by which critical thinking is conducted in communities of inquiry. If confirmed in future studies, this finding would constitute another insight gained through sentence-level analysis by the TAT.

## Conclusion

In developing the practical inquiry model, Garrison et al. (2001) wrote that the fundamental problem was to see and assess thought processes "through the traces of the process that are made visible and public in the transcript" (p. 12). They went on to note that this process was "inevitably inductive and prone to error," due to the subjective judgments necessarily involved. They also acknowledged that the transcript was itself an incomplete and imperfect record of the group's interactions, and consequently of its learnings, since it lacked a record of all the other interactions engaged in by the participants. Perhaps in response to these perceptions, their analytic model appeared to prize simplicity and generalizability, at the expense of accuracy (by Thorngate's principle of compensatory complexity; Thorngate, 1976, cited in Weick, 1979).

Despite problems with interaction analysis as a means of judging the qualities of online learning experiences, use of transcripts in this way remains one of the few methods available to study important social and cognitive aspects in online learning situations. Problems are greater when the focus is on latent projective variables like critical thinking, whose presence must be inferred from other indicators (Rourke, et al., 2001). In such studies, the more indicators incorporated in the analysis the more likely that accurate analytic judgments will be made, as more potentially causal factors are considered in the research process. (This process is termed *overdetermination* by Weick, 1979, p. 37). In this study, the use of the two models, with their different foci and processes, provided a high level of overdetermination, as shown both by the areas of consensus and by the unique contributions made by each.

This paper offers evidence that aspects of the PI model's phases may be usefully elaborated at the level of the sentence by the TAT. In some cases, the greater detail provided by the TAT showed some of the concrete communications and interpersonal strategies (Witte, 1983) on which the phases of the PI model were based (especially in relation to the nature of *triggers,* and the interpersonal and network focus of postings coded *other*). It also appeared that the iterative nature of the PI model, and the conceptual interconnectedness of the model's phases, provide a promising conceptual guide for researchers studying the "sociocognitive process" (Garrison et al., 2001, p. 13) of interaction through CMC. While questions and even equivocalities remain (Garrison, et al., 2001, p. 11), these are not signs of failure, but of the "dilemmas that face those who choose as their topic of interest phenomena that are complex, fluid, collective" (Weick, 1979, pp. 11 –12).

## References


Agreement observer. (2000, October). Chest x-ray. Retrieved March 21, 2001, from http://www.chestx-ray.com/statistics/kappa.html.

Anderson, T., Rourke, L., Garrison, D.R. & Archer, W. (2001). Assessing teaching presence in a computer conferencing context. *Journal of Asynchronous Learning Networks, 5*(2), pp. 1- 17.

Best, J.W. (1970). *Research in education (2nd edition).* Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Blanchette, J. (2001). Questions in the online learning environment. *Journal of Distance Education, 16*(2). Retrieved December 28, 2002, from http://cade.athabascau.ca/vol16.2/blanchette.html

Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H. & Krathwohl, D.R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain.* New York: David McKay Co., Inc.

Borg, W. & Gall, M. (1989). The methods and tools of observational research. In W. Borg & M. Gall (Eds.), *Educational Research: An introduction (5th ed.,* pp. 473 – 530). London: Longman. (Cited in Garrison, D. R., Anderson, T., & Archer, W., 2001).

Bostock, S. (1997). Designing web-based instruction for active learning. In Badrul Khan (Ed.), *Web-based instruction* (pp. 225-230). Englewood Cliffs, NJ: Educational Technology Publications.

Bullen, M. (1998). Participation and critical thinking in on-line university distance education. *Journal of Distance Education, 13*(2), pp. 1-32.

Collison, G., Elbaum, B., Haavind, S. & Tinker, R. (2000). *Facilitating online learning: Effective strategies for moderators.* Madison, Wisc.: Atwood Publishing.

Commonwealth of Learning. (1993). *Student support services: Towards more responsive systems.* Vancouver: Author.

Curtis, D. & Lawson, M. (2001). Exploring collaborative online learning. *Journal of Asynchronous Learning Networks, 5*(1). Downloaded October 15, 2001, from: http://www.aln.org/alnweb/journa;/Vol5_issue1/Curtis/curtis.htm.

Davie, L. & Wells, R. (1992). Empowering the learner through computer-mediated communication. In M.G. Moore (Ed.), *Distance education for corporate and military training (Readings in Distance Education, No. 3*, pp. 104 - 112). The Pennsylvania State University, American Centre for the Study of Distance Education.

Dede, C. (1996). The evolution of distance education: Emerging technologies and distributed learning. *The American Journal of Distance Education, 10*(2), pp. 4 - 36.

Fahy, P. J. (2001). Addressing some common problems in transcript analysis. *IRRODL, 1*(2). Available: http://www.irrodl.org/content/v1.2/research.html#Fahy.

Fahy, P. J. (2002a). Epistolary and expository interaction patterns in a computer conference transcript. *Journal of Distance Education, 17*(1), pp. 20 – 35.

Fahy, P. J. (2002b). Use of linguistic qualifiers and intensifiers in a computer conference. *The American Journal of Distance Education, 16*(1), pp. 5 – 22.

Fahy, P. J. (2003). Indicators of support in online interaction. *International Review of Research in Open and Distance Learning, 4*(1). Available: http://www.irrodl.org/content/v4.1/fahy.html.

Fahy, P. J. (2004). Use of the Bales model for analysis of small group communications in the analysis of interaction in computer-based asynchronous conferences. Paper presented to the IASTED International Conference on Knowledge Sharing And Collaborative Engineering (KSCE 2004), November 22-24, St. Thomas, US Virgin Islands. In Boumedine, M and Ranka, S. (Eds.), *Proceedings of the IASTED International Conference on Knowledge Sharing And Collaborative Engineering* (pp. 44 – 49). Calgary: ACTA Press.

Fahy, P. J., Crawford, G., Ally, M., Cookson, P., Keller, V. and Prosser, F. (2000). "The Development and Testing of a Tool for Analysis of Computer Mediated Conferencing Transcripts." *Alberta Journal of Educational Research, 46*(1), pp. 85 - 88.

Fahy, P. J., Crawford, G. & Ally, M. (2001). Patterns of interaction in a computer conference transcript. *International Review of Research in Open and Distance Learning*, *2*(1). Available: http://www.irrodl.org/content/v2.1/fahy.html.

Fahy, P. J., & Ally, M. Student learning style and asynchronous CMC in distance education. *American Journal of Distance Education* (in press).

Fulford, C. P. & Zhang, S. (1993). Perception of interaction: The critical predictor in distance education. *The American Journal of Distance Education, 7*(3), pp. 8 – 21.

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of distance education, 15*(1), pp. 7 – 23.

Garton, L., Haythornthwaite, C., & Wellman, B. (1997). Studying online social networks. *Journal of Computer-Mediated Communications, 3*(1). Retrieved December 17, 1999, from: http://www.ascusc.org/jcmc/vol3/issue1/garton.html

Gibson, C. C. (1996). Collaborative learning in action via computer mediated conferencing. In *Proceedings of the 12th Annual Conference on Distance Teaching and Learning: Designing for Active Learning* (pp. 121 – 125). Madison: University of Wisconsin.

Gunawardena, C., Lowe, C. & Anderson, T. (1997). Analysis of a global on-line debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research, 17*(4), pp. 395-429.

Haughey, M. & Anderson, T. (1998). *Networked Learning: The pedagogy of the internet.* Montreal: Chenelière/McGraw-Hill.

Henri, F. (1992). Computer conferencing and content analysis. In A. Kaye (Ed.), *Collaborative learning through computer conferencing: The Najaden papers* (pp. 117-136). Berlin: Springer-Verlag.

Herring, S. C. (1996). Two variants of an electronic message schema. In Susan C. Herring (Ed.), *Computer-mediated communication* (pp. 81-106). Philadelphia: John Benjamins Publishing Co..

Jonassen, D., Davidson, M., Collins, M., Campbell, J. & Bannan Haag, B. (1995). Constructivism and computer-mediated communication in distance education. *American Journal of Distance Education*,9(2), pp. 7 – 26.

Jonassen, D. (1998). Designing constructivist learning environments. In C.M. Reigeluth (Ed.), *Instructional Theories and Models, 2nd ed* (pp. 1-21). Mahwah, NJ: Lawrence Erlbaum. Available: http://www.ed.psu.edu/insys/who/jonassen/cle/cle.html.

Kanuka, H. & Anderson, T. (1998). Online social interchange, discord, and knowledge construction. *Journal of Distance Education, 13*(1), pp. 57 – 74.

Keller, V. (1999). *Adaptation and application of a transcript analysis tool to analyze a computer-mediated communication (CMC) distance education course transcript*. Unpublished master's thesis, Athabasca University, Athabasca, Alberta.

Marttunen, M. (1998). Learning of argumentation in face-to-face and e-mail environments. (ERIC Documents: ED 422791.)

Mayer, R. E. (2001). *Multimedia learning.* New York: Cambridge University Press.

McKlin, T., Harmon, S. W., Evans, W. & Jones, M. G. (2002). Cognitive presence in Web-based learning: A content analysis of students' online discussions. ITFORUM Paper #60. Downloaded March 22, 2002, from: http://it.coe.uga.edu/itforum/paper60/paper60.htm.

Poscente, K. (2003). *Trigger analysis in computer-mediated conferencing.* Unpublished master's thesis, Athabasca University, Athabasca, Alberta, Canada.

Puustjärvi, J. (2004). Dynamic ontologies in information retrieval. Proceedings of the IASTED International Conference on Knowledge Sharing and Collaborative Engineering, St. Thomas, USVI, November 22 – 24, pp. 115 – 120.

Riffe, D., Lacy, S., & Fico, F. (1998). Analyzing media messages: Using quantitative content analysis in research. Mawah, N.J.: Erlbaum. (Cited in Rourke, et al., 2001.)

Roblyer, M. D. & Schwier, R. A. (2003). *Integrating educational technology into teaching, Integrating educational technology into teaching, Canadian edition.* Toronto: Pearson Education Canada Inc.

Romiszowski, A. (1997). Web-based distance learning and teaching: Revolutionary invention or reaction to necessity? In Badrul Khan (Ed.), *Web-based instruction* (pp. 25 - 37). Englewood Cliffs, NJ: Educational Technology Publications.

Rose, M. A. (2004). Comparing productive online dialog in two groups styles: Cooperative and collaborative. *The American Journal of Distance Education, 18*(2), pp. 73 -- 88.

Rourke, L., Anderson, T., Garrison, D. R. & Archer, W. (1999). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education, 14*(2), pp. 50 - 71.

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (2001). Methodological issues in the content analysis of computer conference transcripts. *International Journal Of Artificial Intelligence and Education, 12*. Available from: http://cbl.leeds.ac.uk/ijaied/.

Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development, 52*(1), pp. 5 - 18.

Schrage, M. (2003). Flaming ideas. *Technology Review, 106*(2), p. 23.

Simon, J. & Burstein, P. (1985). *Basic research methods in social science* (3rd Edition). Toronto: McGraw-Hill, Inc.

Strother, J. (2002). An assessment of the effectiveness of e-learning in corporate training programs. *International Review of Research in Open and Distance Learning, 3*(1). Downloaded March 11, 2003, from http://www.irrodl.org/content/v3.1/strother.html.

Thorngate, W. 1976. "In general" vs. "it depends": Some comments on the Gergen-Schlenker debate. *Personality and Social Psychology Bulletin 2,* pp. 404 – 410. (Cited in Weick, 1979).

University of Colorado. (1999). Calculating Cohen's *Kappa*. Retrieved March 21, 2001: [http://www.Colorado.EDU/geography/gcraft/notes/manerror/html/kappa.html

Walther, J. B. (1996). Computer- mediated communication: Impersonal, interpersonal and hyperpersonal interaction. *Communication Research, 20*(1), pp. 3 – 43.

Weick, K. E. (1979). *The social psychology of organizing, 2nd edition.* Reading, Mass.: Addison-Wesley Pub. Co.

Witte, S. (1983). Topic structure and revision: an exploratory study. *College Composition and Communications (34)*, 3, pp. 313 - 341.

Yates, S. J. (1997). Gender, identity and CMC. *Journal of Computer-Assisted Learning, 13*, pp. 281 – 290.

Zhu, E. (1996). Meaning negotiation, knowledge construction, and mentoring in a distance learning course. In *Proceedings of Selected Research and Development Presentations at the 1996 National Convention of the Association for Educational Communications and Technology* (18th, Indianapolis, IN). (ERIC Documents: ED397849.)

## About the Author

**Patrick J. Fahy** is an Associate Professor in the Centre for Distance Education (CDE), Athabasca University. His career has included teaching and research at all levels of adult education, private sector management and training experience, and consulting with government and private sector clients. Presently, in addition to developing and teaching educational technology courses in the CDE, Pat coordinates the Advanced Graduate Diploma in Distance Education (Technology) program, and the CDE's annual Athabasca University Distance Education Technology Symposium (DETS). His research interests include online interaction media, and CMC transcript analysis. He is a former president of the Alberta Distance Education and Training Association (ADETA). He can be contacted as follows:

Patrick J. Fahy, Ph.D., Associate Professor
Centre for Distance Education
Athabasca University
1 University Drive
Athabasca, Alberta, Canada T9S 3A3

Phone: 866-514-6234 E-mail: patf@athabascau.ca

**Editor's Note**: Courses with strong visual, interactive, and decision making components make excellent online courses. This paper shares five years of experience in adapting a course – and online software – for a statistics course in economics. Its strengths as perceived by students are interactivity and online resources. Additional assistance and activities would further strengthen this course.

# Learning Statistics in a Shared Virtual Campus. Summarizing a Five-Year Experience

**Ana Jesús López and Rigoberto Pérez**

## Introduction

The development of Information and Communication Technologies (ICT) has introduced substantial social changes, including the dissemination of statistical information and teaching and learning methods. As a consequence, statistical courses can benefit greatly from the increasing availability of information, the use of statistical packages and the possibilities of Internet as a teaching tool.

In this paper we describe the on-line course "Economic Data Analysis" which is offered in the virtual campus of the University of Oviedo, AulaNet and also in the so-called G9 Shared Virtual Campus. The work contains a brief presentation of the learning method, describing the main teaching tools and the educational software ADE+. A "facts and figures" section is also included, summarizing our five-year experience with this on-line subject.

The paper ends with some concluding remarks and a list of bibliographical references.

***Keywords:*** Statistics, Internet, ICT, e-Learning, AulaNet, G9 Shared Virtual Campus, Economic Data Analysis, educational software, ADE+, online polls, European Space for Higher Education.

## 1. Statistics in the Information Society

Information and Communication Technologies (ICT) have deeply affected the production, dissemination and treatment of statistical information. The Statistical Offices have played a fundamental role in this new situation, since their Internet servers provide fast access to vast amounts of information and resources thus becoming an increasingly popular way of finding statistical information.

In 1929 H.G. Wells predicted:

> *"The time may not be very remote when it will be understood that for complete initiation as an efficient citizen of one of the new great complex world wide states that are now developing, it is as necessary to be able to compute, to think in averages and maxima and minima, as it is now to be able to read and to write."*

Today most people would agree with this thought since a certain statistical knowledge is required for the correct interpretation of daily economic information, such as the Consumer Price Index or the employment data. Furthermore, most students are bound to deal with statistical information in their professional future.

Teaching and learning methods have experienced substantial changes during the last years, while printed materials have progressively been complemented by audios, videos, computers and -more recently- digital technology.

In this gradual process Internet has introduced a wide variety of pedagogical resources, including the access to on-line teaching material (presentations, working papers, examples, links of interest, suggested activities, …) and also communication tools (e-mail, chats, forums, distribution lists, video-conferences,...). The efficient use of these facilities would introduce some outstanding advantages, including not only space and time flexibility but also increasing student interaction in the learning process.

The term E-learning means, according to the European Commission, "*using new multimedia technologies and the Internet to improve the quality of learning.”* In this context, teachers must face the challenge of using the ICT facilities for the development of high-quality educational materials, also improving the level of communication and satisfaction of teachers and students.

This paper describes our experience with the on-line course “Economic Data Analysis” (ADEnet), included in the virtual campus of the University of Oviedo and also in the Shared Virtual Campus developed by the G9 Group of universities.

Section 2 provides an overview of the course and its learning resources, while section 3 briefly describes the main characteristics of our software ADE+. The description of tutorials, communication facilities and evaluation is presented in section 4. Next, section 5 summarizes some of the most outstanding facts of our five-year experience, including information about student’s profiles, results and opinions. The paper ends with concluding remarks in Section 6.

## 2. Economic Data Analysis: Learning Resources

“Economic Data Analysis” is a practice-oriented subject, which has been designed considering the students as users of the statistical information and trying to make an efficient use of Information and Communication Technologies (ICT), mainly the Internet. This optional subject was launched in the course 1999/2000 and is currently offered in the Shared Virtual Campus developed by the G9 Group, which includes the universities of Cantabria, Castilla-La Mancha, Extremadura, Illes Balears, La Rioja, Pública de Navarra, Oviedo, País Vasco and Zaragoza.

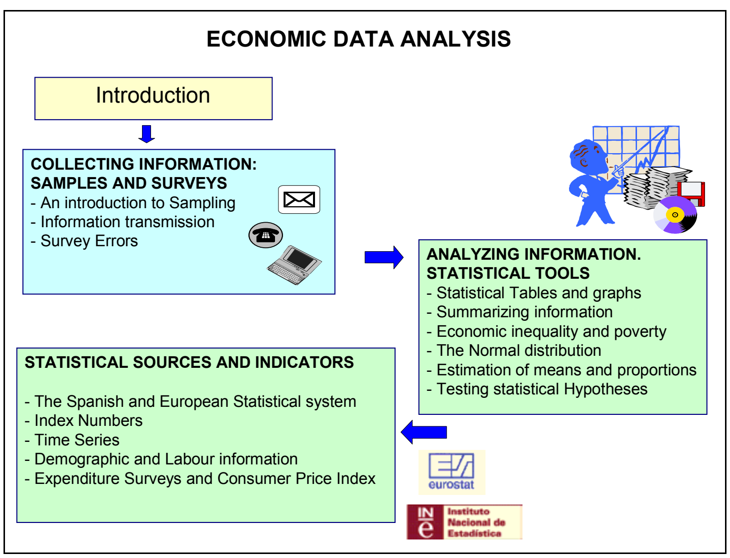



**Figure 1. Economic Data Analysis - Course Syllabus**

As figure 1 shows, the syllabus of this on-line course contains a three module classification, following the sequence of the origin, treatment and dissemination of statistics.

The first approach to each item is provided by a "virtual lesson", designed as a twelve-minute multimedia presentation, offering an overview of its main contents. Afterwards students can access a wide variety of learning resources, including electronic books, dynamic presentations, interactive questions, figures, links, glossaries of terms, web references, self-assessments … The practical contents are based on the software ADE+, which is described in the next section.

The statistical websites play an important role in the course since they clearly show the ways of access and treatment of the economic information. Students are encouraged to become familiar with the main statistical servers, such as the Spanish National Statistical Institute, (http://www.ine.es), the European Office for Statistics Eurostat (http://europa.eu.int/comm/eurostat), United Nations (http://www.un.org) or the World Bank (http://www.worldbank.org).

With the aim of encouraging the participation of students, the course includes several interactive questions, which are easily corrected from the web allowing the students to check their level of knowledge. The critical perspective is emphasized through the inclusion of proposed questions. A "Statistical Mistakes" section is provided, showing some historical anecdotes and risks of statistical analysis.

As an example, let us consider the lesson studying the Inequality and Poverty Measurement. An easy introduction to this field is presented through the story of a wealthy family and its legacy distribution, analyzing the inequality degree related to different options.

The main concepts of this chapter (income and population shares, Lorenz curve, Gini index, inequality measures, poverty line, ...) are introduced with graphical assistance, also including many links to statistical websites as the United Nations, the World Bank or the Spanish National Statistical Institute (INE). In order to encourage their participation, students are asked about the expected impact on inequality of some distributive policies such as proportional taxes or income transfers, whose effects are graphically displayed through animated presentations.

The chapter ends with some considerations about poverty, including a brief report on the 1998 Nobel laureate in Economics, Amartya Sen, summarizing his contributions to the measurement of inequality and poverty.

Once the lesson is finished, a self-assessment is suggested. In this option a battery of questions is randomly selected and corrected from the web, the student marks being transferred to the subject's database in AulaNet.

## 3. The Software Ade+

Following the "learning by doing" approach, the practical contents of "Economic Data Analysis" are based on the software ADE+, specifically developed for teaching purposes. This application has been developed and registered by R. Pérez and A.J. López (license 1996/33/27694 of the Intelectual Property Provincial Register of Asturias) and is available from the AulaNet website: http://www.aulanet.uniovi.es/ade+/.

Although there is a wide variety of available statistical packages, suitable software for teaching purposes is not always easy to find, especially for introductory courses.

On the other hand, the use of professional applications does not seem adequate since most packages become self-sufficient, leading users to adopt a passive role and restricting their learning potential. One solution commonly adopted to promote participation of students in the practical sessions of Statistics is the use of spreadsheets. According to our experience, this option has some advantages:

- Spreadsheets are user-friendly tools
- Many statistical offices and institutions provide excel workfiles which can be easily downloaded from their Internet servers
- Students are allowed to gradually build the required operations
- Results can be automatically updated

Nevertheless, we have also detected some important objections related to the educational use of the spreadsheets:

- Statistical functions and commands are not always correctly defined
- Spreadsheets do not usually provide options which are commonly used in economic statistics, such as index numbers or rates of growth.

These disadvantages have suggested the development of didactical software which is usually included in some statistics textbooks and/or available in Internet. In fact this was also the motivation behind the software ADE+ which aims to support interactive data analysis, allowing the students to build and analyze their data sets.

As it is shown in figure 2, ADE+ is organized in three different areas: a text editor, a spreadsheet-like data table and an object container. This structure allows a comprehensive treatment of the statistical information, including data collection, graphical representations, statistical analysis and interpretation of results.

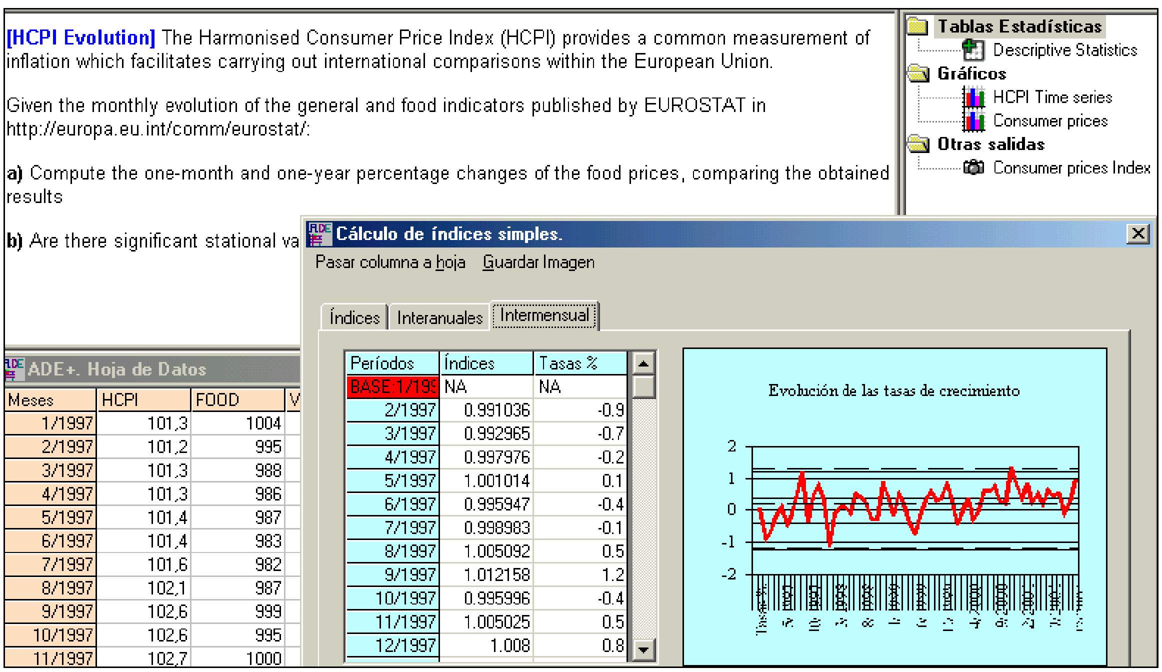



**Figure 2. The ADE+ Software**

A wide variety of statistical options is available from the menu, including sampling, data tabulation and representation, descriptive statistics, joint analysis, index numbers, time series, probability distributions and statistical inference.

As Pérez & López (2003) state, the use of the ADE+ software has proved to be satisfactory in both presential and virtual education. In the case of "Economic Data Analysis" the students can access a wide variety of workfiles with solved and proposed questions and also an online user guide explaining all the available facilities.

## 4. Tutorials, Communication and Evaluation

Although there is a general agreement about the e-learning potential, some risks have also been detected, mainly referring to the isolation of students. Therefore, the success of an on-line course depends to a great extent on its communication and evaluation facilities.

Some recent works find that the dialogue is more difficult when it is mediated by ICT, but at the same time the amount of dialogue can be greater in e-learning courses than in traditional ones. In the case of our course "Economic Data Analysis" dialogue is strongly encouraged and therefore tutorials are provided in several ways, including e-mail, chats, forums and video-conferences.

On-line assistance connects students with their tutors, providing quick answers (within 24 hours) to their questions and comments through e-mail. This option is widely used by students together with a shared board conceived as a permanent forum. Both tools are based on asynchronous communication, allowing flexible participation of students any time they have a question.

Furthermore, students have access to a weekly chat, allowing synchronous discussion with classmates under the coordination of a teacher. For those preferring a face-to-face debate, video-conference tutorials are also available.

According to our experience, students attach great value to having access to personalized information. Therefore, confidential reports are available, providing indicators of the student's progress (such as assessment results) and specific comments about their personal work.

The evaluation system of "Economic Data Analysis" has been designed trying to achieve coherence with the learning process. More specifically, the final grades are obtained according to the following criteria:

Self-assessment results (30%)

- Marks of the personal work sent to tutors (30%)
- Marks of the final on-line exam (30%)
- Activity and participation Indicators (10%)

It must be noticed that the online evaluation is an outstanding characteristic of this subject, providing a distinctive feature from most virtual experiences.

Online evaluation of Economic Data Analysis is designed as a practical exam allowing students to show their ability in the use of statistical tools and the interpretation of economic indicators. A key element in this evaluation is ADE+, since students are requested to solve some problems with this software, the workfiles being self-stored and forwarded to the AulaNet platform.

As expected, the implementation of the online evaluation required special efforts in the context of the G9 Shared Virtual Campus, whose members have developed a coordinated strategy paying special attention to questions such as timetables, security conditions, student isolation or tutor accessibility.

Since students should have full academic recognition of the credits coursed at each of the universities included in the G9 Group, the success of the shared virtual campus needs great attention being paid to both technical and administrative aspects.

## 5. Facts and Figures

Our five-year experience with the subject “Economic Data Analysis” has provided some useful information about academic results and students´ opinions.

As can be seen in Figure 3, the number of students shows an increasing path both for the University of Oviedo and the other institutions included in the G9 Shared Virtual Campus.

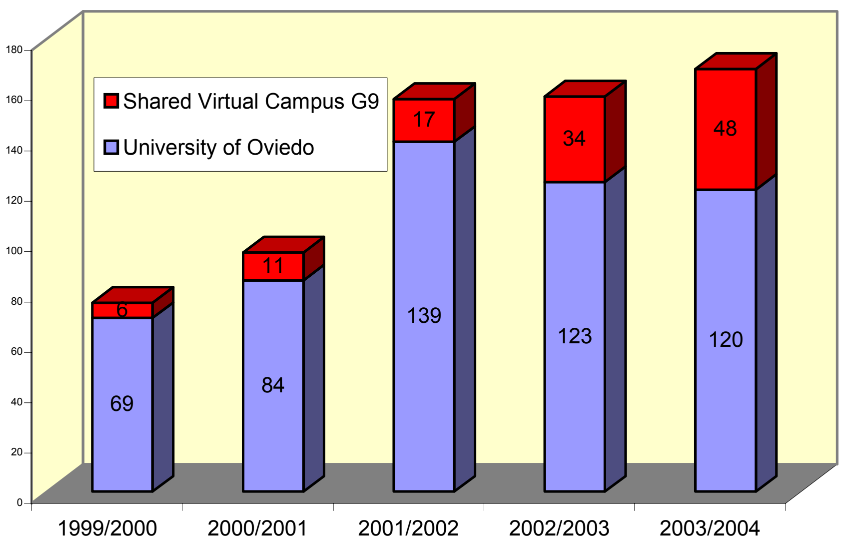



**Figure 3. Evolution of Students in “Economic Data Analysis”**

It must be stressed that the positive evolution of this subject refers not only to the increasing number of enrolled students but also, and most important, to their rate of participation in the course, which could be interpreted as the complementary of the percentage of desertion.

This is quite an interesting fact, since desertion has been identified as one of the most outstanding risks of the e-learning experiences. Although this fact could be the consequence of different factors (including the lack of information, technical difficulties or personal reasons) we have noticed that some of these problems can be avoided by making an extra effort to inform, to help and to encourage students in their learning process.

In order to collect the opinions of all the agents involved in the e-learning process, the coordinators of the G9 shared virtual campus have developed a wide variety of tools, including on-line surveys for students and teachers and also specific interviews for those students who have given up their virtual experiences. Furthermore, the annual meetings of the G9 Shared Virtual Campus provide teachers, students and administrators with an excellent opportunity for sharing their experiences.

According to the online surveys, the students of the G9 Shared Virtual Campus are mainly motivated by the use of new learning methods and their flexibility. As Salinas et al (2002) summarize, these students are connected to the Shared Virtual Campus an average of three hours per week, and most of them find that the university infrastructures do not satisfy the e-learning requirements.

According to their answers, the students find the learning materials as adequate or very adequate (82.6%) of a high quality and well structured (80.4%) and available in time (71.7%). Figure 4 represents the main results referring to the course “Economic Data Analysis” and to the G9 Shared Virtual Campus as a whole.

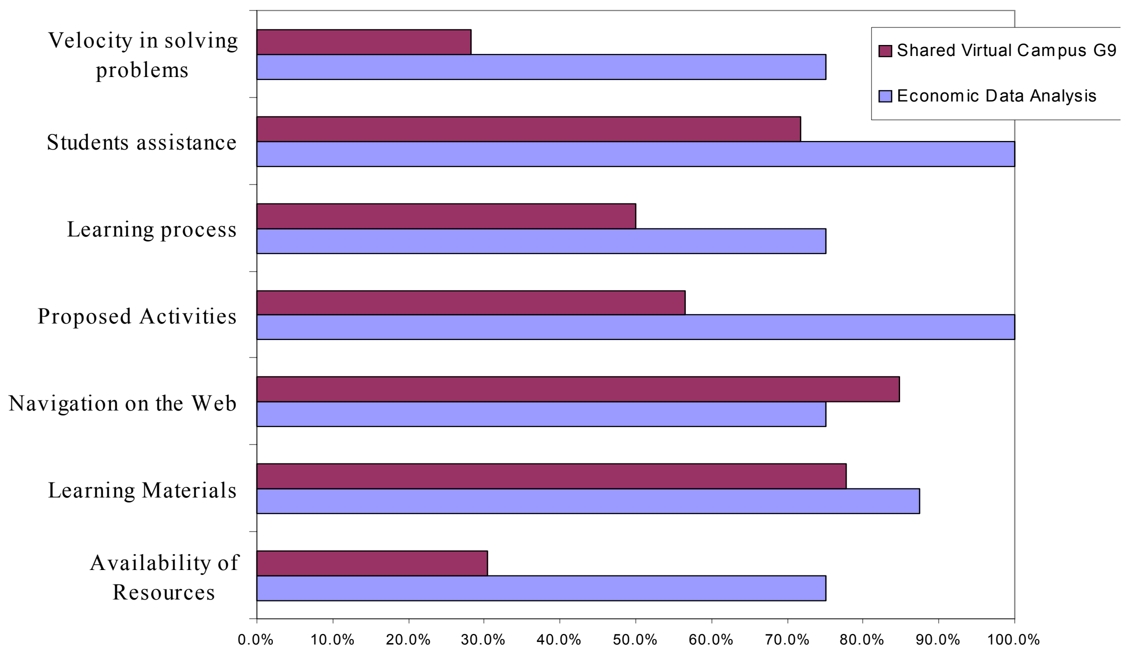



**Figure 4. Student Online surveys: Proportion of answers “Very Satisfactory” or “Satisfactory”**

In the case of our course “Economic Data Analysis” students have repeatedly been interviewed about their experience, including technical questions, pedagogical aspects and specific educational tools. In general terms, the experience has been positively evaluated and the main results are summarized in Figure 5, where the average score assigned to each item (in a scale from 0 to 10) is represented together with the Pearson´s coefficient of variation.

| | Average score | Pearson´s coefficient of variation |
|---|---|---|
| Virtual Lessons | 6.73 | 29.2% |
| Practical contents | 8.22 | 14.4% |
| Software ADE+ | 7.51 | 20.5% |
| Self-evaluation | 8.27 | 19.5% |
| Interactive questions | 7.20 | 24.5% |
| Tutorials | 7.81 | 30.1% |
| Communication facilites (forum, chat, …) | 4.14 | 72.8% |

**Figure 5. "Economic Data Analysis":**
**Results of Students online surveys**

In general terms, the best scores are related to the most participative tools, such as self-assessments, interactive questions or practical contents solved with ADE+.

Although the students also appreciate the personal assistance and tutorials, the use of communication tools (chat, forum, video-conference, ...) has obtained lower average scores with high relative dispersion. This lack of representativeness shows the existence of two different groups of students with active and passive behaviours.

Technical difficulties and low speed or the network remain as the most negative aspects, although a considerable improvement has been observed during the period.

Finally, it is important to stress that once the students have became familiar with these new learning methods and experienced their advantages, most of them are highly interested in further e-learning experiences.

## 6. Concluding Remarks

The development of the Information Society has led us to a new educational context, in which students can learn statistics in a more realistic, flexible and participative way. Our five-year experience with the online subject Economic Data Analysis has been satisfactory for both students and teachers, showing the didactical power of Internet and allowing the implementation of multimedia educational material (MEM), links to statistical servers, self-assessments and an on-line examination.

The use of the statistical software ADE+, specifically developed for teaching purposes, has shown many advantages since it allows students to easily access and analyze economic information.

The described experience could be helpful in the design of new e-learning courses, based on multimedia technology and pedagogy-oriented. Furthermore, since e-learning stimulates exchanges and collaborations, the experience of the G9 Shared Virtual Campus could be successfully extended to other university groupings established in different spatial areas.

In the European context these strategies could contribute to the achievement of the Lisbon Summit (2000) strategic goal: *"to make out of the European Union the world's most competitive and dynamic knowledge-based economy, capable of sustainable economic growth and with more and better jobs and greater social cohesion"*. In fact, the e-Learning initiative launched by the

European Commission under the title "*Designing tomorrow's education*" aims to create the appropriate conditions for the development of contents, services and learning environments, improving cooperation and dialogue at regional, national and European levels and between all the participants in the field (universities, schools, decision-makers, administrations, ...).

According to a recent study by PLS Ramboll (2004), most European universities show a positive attitude towards ICT integration and e-learning, although the support and priority allocated by university management is a critical obstacle in many of them.

This research distinguishes four university clusters, respectively defined as the front-runners (18%), the co-operating universities (33%), the self-sufficient (36%) and the sceptical (15%), concluding that most universities still face a severe challenge in terms of incorporating the results and experiences gained from development projects into their overall strategy.

Since we are now facing the Bologna process, with the aim of establishing a European Higher Education Area, in which staff and students can move with ease and have fair recognition of their qualifications, it seems quite clear that Information and Communication Technologies could be extremely helpful for the achievement of these goals. Therefore, we think it is time to share experiences, stimulate exchanges, facilitate interaction and collaboration and recognize our students as competent, active and constructive partners in higher education.

## References

Behar, R. & Grima, P. (2001). Mil y una dimensiones del aprendizaje de la estadística. *Estadística Española*, 43, 148, 189-207.

Belli, G. (2000). The teaching of statistical consulting skills. *Paper presented at the Meetings of the American Educational Research Association*, New Orleans, April 24-28.

Ben-Zvi, D. (2001). Technological Tools in Statistical Education, In Jornades europees d'estadística. L'ensenyament i la difusió de l'estadística. Ed. Conselleria d'Economia, Comerç i Indústria, Palma de Mallorca. 207-220.

Bottino, R.M. (2001). Advanced Learning Environments, In *Computers and Education: Towards an Interconnected Society*, M. Ortega and J. Bravo Ed., Kluwer Academic Publishers, 11-27.

Chanter, D.O. (2001). Some Anecdotes and Analogies for Illustrating Statistical Ideas, in *The best of Teaching Statistics*, http://science.ntu.ac.uk/rsscse/ts/bts/chanter/text.html

European Commission (2002). *eEurope Benchmarking Report,* Brussels. http://europa.eu.int/information_society/eeurope/benchmarking/index_en.htm

European Commission (2004). http://elearningeuropa.info

Gärdin, O. (2002): The New Economy New challenges for the statistical system, In *The International Association for Official Statisticians (IAOS) Conference*, London.

López, A.J. & Pérez, R. (2001a). An experience on virtual teaching: AulaNet, in *Computers and Education: Towards an Interconnected Society*, M. Ortega and J. Bravo Ed., Kluwer Academic Publishers, 207-214.

López, A.J. & Pérez, R. (2001b). La enseñanza de Estadística con nuevas tecnologías. Algunas experiencias universitarias, in *Jornades europees d'estadística. L'ensenyament i la difusió de l'estadística*. Ed. Conselleria d'Economia, Comerç i Indústria, Palma de Mallorca, 221-232.

Mills, J.D. (2002). Using Computer Simulation Methods to Teach Statistics: A Review of the Literature, *Journal of Statistics Education*, 10 (1), http://www.amstat.org/publications/jse/v10n1/mills.html

Moore, D. (1997). New pedagogy and new content: the case of statistics. *International Statistical Review 65*, 123-165.

Mulekar, M. (2000). Internet resources for AP Statistics Teachers, *Journal of Statistics Education*, 8 (2), http://www.amstat.org/publications/jse/secure/v8n2/mulekar.cfm

Pérez, R. & López, A.J. (2003). Economic Data Analysis with ADE+. Summarizing presential and virtual experiences, *Proceedings of the IASE Satellite Conference on Statistics Education and the Internet*, http://www.ph-ludwigsburg.de/iase/proceedings/index.html

Pérez, R., López, A.J. & Río M.J. (2002). Learning statistics online. Facts and figures from a three-year experience, *International Conference on Information and Communication Technologies in Education*, vol. 3, 1441-1444, Badajoz.

Pérez, R. et al (1997). *Análisis de datos económicos I. Métodos descriptivos*. Ed. Pirámide, Madrid.

PLS Ramboll (2004). *Studies in the context of the E-Learning initiative: Virtual Models of European Universities*, EU Commission, DG Education and Culture.

Reading, C. (1998). Reactions to Data: Students' Understanding of Data Interpretation. In L. Pereira-Mendoza, L.S. Kea, T.W. Kee & W.-F. Wong (Eds), *Statistical Education – Expanding the Network. Proceedings of the Fifth International Conference on Teaching of Statistics.* (1427-1433). Singapore.

Reding, V. (2001). The role of the European Community in creating the knowledge and information society, *Zentrum für Europäischen Integrationsforschung, University of Bonn,* http://www.europa.eu.int/comm/elearning

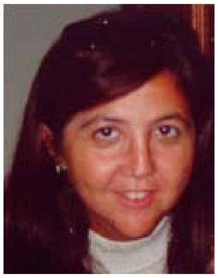
Relander, T. & Kunnas, R. (1999). The Impact of the Internet on the Production and Dissemination of Statistics, in *Meeting on the Management of Statistical Information Technology*, Statistical Commission and Economic Commission for Europe.

Salinas et al. (2002). Red de servicios y estructuras comunes en el Campus Virtual Compartido del G9 de Universidades, *Informe final del proyecto EA2002-0109, Programa de Estudios y Análisis*, Secretaría de Estado de Educación y Universidades.

Tiffin, J.and Rajasingham, L. (1995), *In Search of the Virtual Class: Education in an Information Society*, London and New York.

Yilmaz, M. (1996). The Challenge of Teaching Statistics to Non-Specialists, *Journal of Statistics Education*, 4 (1), http://www.amstat.org/publications/jse/v4n1/yilmaz.html

## About the Authors



**López, Ana Jesús**
Department of Applied Economics
University of Oviedo
Campus del Cristo, s/n
33006 Oviedo (Spain)

Tel: 34 985103759
anaj@uniovi.es

**Ana Jesús López** is Professor of Statistics and Econometrics at the University of Oviedo, Spain. She received her Ph.D. in Economics at the University of Oviedo (1991) and Bachelor in Economic Sciences at the same University (1986).

Her research activities are related to three fields: measurement of economic inequality, regional modelling and forecasting, and the socio-economic impact of the Information and Communication Technologies (ICTs), mainly the e-learning.

Dr. Lopez has been involved in several research projects and she is author of several publications in international and Spanish journals.

**Pérez, Rigoberto**
Department of Applied Economics
University of Oviedo
Campus del Cristo, s/n
33006 Oviedo (Spain)

Tel: 34 985103748
Mail: rigo@uniovi.es

**Rigoberto Pérez** is Professor of Statistics and Econometrics at the University of Oviedo, Spain. He is Ph.D. in Economics at the University of Oviedo (1985) and Bachelor in Mathematics at the University of Santiago de Compostela (1978).

Professor Pérez published several books and articles in international and Spanish journals, has been involved in several research projects, and has supervised seven Doctoral Dissertations and eight Post-Graduate Dissertations. In his current position of Director of the Innovation Center of the University of Oviedo and the Shared Virtual Campus of the G9 Group (including nine Spanish Universities) he is also involved in several projects dealing with the ICTs impact, the e-learning and the development of suitable quality indicators.

**Editor's Note**: At the beginning of my teaching career I had the good fortune to watch a tutor teach lip-reading to 3 year old deaf child at the John Tracy Clinic for the Deaf at University of Southern California. At the end of the session, the teacher came into our viewing room behind a one-way glass mirror. Her colleague immediately went through a list of suggestions to make the lesson more effective. I was shocked. I thought the teacher had done a brilliant job. When the teacher left I questioned the critical comments. "We do this for each other at every session. It is the only way we can learn; the only way we can improve." At that moment I realized that mutual trust and a common goal opened the way for continuous improvement of a life-changing process for these children. The process brought out the best thinking of teachers-helping-teachers, resulting in outstanding performance. That is what this article is about.

Don Perrin, Executive Editor

# Flying or Falling:
## Benefits and Pitfalls of Online Peer Review Programs in Distance Education

**Kim Blum and Brent Muirhead**

## Introduction

One author of this article had a rare opportunity to go river rafting in Australia. On the second trip, the river guiding company placed the author in a raft with the author's grown son and six members of an All Aussie Rules over-30 football team. An Aussie All Rules Football game is played with few fixed rules, no padding, and men over seven feet tall that weight at least 250 pounds of muscle. They have massive scars and scrapes from playing the game.

The author felt, as a result of a previous successful rafting trip where rapids were very difficult, this lower-classed river grade should be quite easy. Rivers for rafting are graded by difficulty, one being the easiest and six almost impossible to navigate by a paddle and raft. The author felt that previous experience on a class six river ensured that, because of this river's easier classification, the raft would *fly* down the river in an enjoyable manner. The author was wrong.

After boarding the raft, the river guide instructed his team on how to use the paddle, lean right, lean left (to avoid rocks and tilt the boat), forward paddle, and back paddle. The worst-case instruction was given and all rafting team members practiced this successfully in the entrance's calm pool waters – get down and hold on (to the rope), lift the paddle up out of the water. When the guide issues the instruction to *hold on*, rafters are typically afraid because *hold on* means the upcoming rapids are very difficult and wild. After the initial training, the team felt ready to face the turbulent waters of the first rapids. The author's previous experience added confidence. The raft was *flying along the river headed towards the rapids* and confidence levels were high.

As the raft approached the rapids, excitement and cheering erupted from paddlers of the All Aussie Rules Football team. All rafting paddlers closely followed the guide's instructions except the author, who promptly *fell* out of the raft and into the swirling rapids, *failing to succeed in spite of previous training,* and in error about level of ability.

Similar to the author's rafting experience, a successful online peer review is received by online faculty as a wonderful flying feeling of success comparable to the feeling of teamwork and exhilaration of a successful manned raft. An online peer review program can successfully further develop faculty after initial training (Carr, 2005). Unfortunately, an online peer review can also resemble falling rafters as resentful faculty members receive an *evaluation* instead of coaching on online best practices.

This article discusses the pitfalls and benefits of online peer review programs, sharing experiences from administration in higher education, and comparing administration experiences with an online faculty member who has been through the online peer review process. Implications of Peer Review Programs for higher education online faculty and administration are included.

## Flying –Purpose of a Successful Online Peer Review Program

According to the U.S. Department of Education (1999), one in three higher education institutions in the United States offered some type of distance education during the 1997-98 academic years. Nearly 80% of all 4-year and almost two thirds of 2-year public institutions made distance education available to students during this period. Of the institutions that did not offer distance education, 20% planned to offer some type of remote delivery service by 2002 (U.S. Department of Education, 1999). Ross and Klug (1999) found that one of the most effective things an institution can do to promote greater receptivity and support for distance education is to enhance faculty knowledge of distance education best practices. Peer review programs are one of the tools designed with the goal of helping and coaching faculty with best practices for success after initial online faculty training (Peer Monitoring of TRIO Programs, 2005). This is typically implemented when the faculty member has taught at least one online course.

A successful online peer review program rests on the concept of equal-in-status faculty members coaching other faculty members on best practices, giving tips on how to handle online discussions, assignments, feedback and materials, and communicating to the peer reviewee with a positive tone in a collaborative manner. Adams (2002) suggested that the best online faculty support systems are those where online faculty participate in a collaborative team, Faculty in the study expressed a strong degree of satisfaction with support provided by the collaborative model.

At University of Phoenix, Peer Reviewers in the School of Advanced Studies (SAS) observe the online class and provide tips to the faculty member written in a positive tone. Praises are included with all tips, based on a guideline of best practices proven effective online. A *successful* online peer review program should result in faculty incorporating suggestions and feeling that the results are not punitive in nature. Proponents maintain that peer assistance and reviews will help floundering teachers and possibly save their jobs (Pushing for Peer Reviews, 2005). "Opponents argue that it will pit one teacher against another and threaten the unity of local union associations" (para 10). The faculty member should feel that despite the rocks and bumps in the river, the river guide – the reviewer – gave good instructions to keep from falling. In the case of the author, the guide hauled the rafter back in by the life jacket, and proceeded to re-coach the rafter on how to stay in the raft so that the raft could continue to fly with success down the river rapids. A successful online peer review should *help faculty fly instead of fall*.

## Man the Paddles – The Processes of Online Peer Reviews

All faculty in SAS have yearly Online Peer Reviews. Newer faculty are reviewed after three months in order to provide more tips on best practices at an earlier stage. The faculty member receives notification of the review in the last week of an 8-week course, student feedback for one successful student and one struggling student is requested, and the reviewer observes the online class newsgroups and feedback. The trained peer reviewer observes the newsgroups in an unbiased manner, using a checklist of best practices as a guide for key areas including materials, discussions, classroom management, tone, and feedback. For example, SAS has found that online students achieve maximum student learning following the adult as an active leaner model (Knowles, 1984), when the professor facilitates discussion by relating work experiences, theories, and asks questions to stimulate higher levels of Bloom, Bertram and Krathwohl (1964) levels of critical thinking from application to synthesis and evaluation.

The online peer reviewer at SAS reports on the findings and incorporates praise and suggestions. The review is sent anonymously by a processor (the reviewer's name is not released) to the faculty member for review and signature. Follow-up procedures include coaching, addressing questions, and explaining processes and purposes if the faculty member is unclear or has some trepidation about the review.

## Fly Instead of Fall: Pitfalls to Avoid in Online Peer Review Programs

At the School of Advanced Studies (SAS), the underling mission of the online doctoral peer review program ensures that subsequent evaluative assessments by administration, triggered by results of a peer review, do not occur *as an evaluation*. It is critical that administration does not use peer review for any purpose other than faculty development. Online faculty fear that evaluations can result in punitive scheduling or pay reductions. The word *evaluation* is not consistent with goals of the peer review program. Its core mission is faculty-helping-faculty.

SAS determined that administration must avoid acting on information presented in peer reviews. The goal is to help faculty succeed and findings should be highly confidential. A distinct and separate faculty evaluation process must be clearly differentiated from the Online Peer Review program. The Faculty Evaluation process is based on findings of the 1981 Teacher Peer Review program that successfully developed teachers (Pushing for Peer Reviews, 2005) positing that peer reviews can be positive developmental experiences if used in this manner.

For example, one author of this article participated in an additional online peer review program. One hundred and sixty-five faculty complaints resulted from peer reviews because of the evaluative nature of the comments recorded by the peer reviewers. In spite of reassurance from administration, follow-up administrative actions stemmed from information noted in the reviews. The cycle of mistrust escalated until peer review processes and communications were changed.

The opposite occurred in SAS. After two years of online peer reviews, SAS has not received any faculty complaints. Comparing patterns in responses and plans, the author concluded that there were several reasons why one online peer review program succeeded where another failed. The successful peer review program had key differences not found in the unsuccessful program:

- Invitations to become SAS Online Peer Reviewers were the result of *months* of research on all SAS faculty. Only the best of the best faculty were invited based on SEOCS, faculty evaluations and observing many of their online classes.
- Extensive training of Online Peer Reviewers in a formal training workshop presented the purpose and guidelines as well as additional time to practice reviews. Follow-up trainee responses received individual coaching on tone, how to avoid certain words that could be perceived as negative, and tips on how to formulate coaching to ensure a positive review. As part of the training, instruction and practice focused on the sandwich method of praising, tips, and ending in praise as an effective online coaching method. This follows finding by Wolf (2003) that the choice of words in a peer review is critical.
- Peer Reviewers conducted one real review, the trainer edited the review, and peer reviewers received additional coaching. A decision to continue with the peer reviewer or decline to award additional reviews depended on the outcome of the analysis.
- Establishment and maintenance of a Peer Review lounge and a questions contact person.
- SAS faculty were continually reminded that the purpose of the Peer Review Program is faculty helping faculty. This is not an evaluation. Peer reviews are never punitive in nature, and scheduling or pay is never affected. Reminders came to all faculty from *top leadership in SAS* on a frequent basis.

- Faculty could see, after an extended period, that administration truly modeled the purpose of the peer review with faculty coaching faculty with no punitive outcomes., faculty were never contacted after any review and coached as an evaluation and scheduling or pay was not affected. Reductions in the levels of faculty fear resulted, and the grapevine did the rest as faculty relaxed and accepted coaching in reviews.

## Faculty Perceptive of Online Peer Reviews:
### Benefits (Flying instead of Falling)

Reflective online faculty have a positive and visionary perspective on professional development. Educators realize that teaching and learning is an evolving process which requires constant attention, experimenting with various instructional strategies and investigation to acquire more effective methods (Brookfield, 1995). Today's online instructors should strive to be life-long learners who realize that it takes time and diligent study and practice to become an expert. Fear of falling must be ignored as the online educator receives tips from the reviewer – the guide.

Cognitive psychologists stress that it often takes ten years for a person to become an expert (Anderson, 2005; Schacter, 1996). Experts have " …a highly refined and powerful form of elaborate encoding that enables experts to pick out key information efficiently and to imbue it with meaning by integrating it with preexisting knowledge" (Schacter 1996, p. 49). Experts possess two kinds of expertise: routine and adaptive. Routine expertise enables the individual to do problem solving in an effective and timely manner. Adaptive expertise skills are those which help people to develop strategies that fit the particular situation (Eysenck, 2001).

Online distance faculty should consider peer reviews as an intentional way to cultivate their expertise. Teaching should be considered a craft and the word *craft* highlights that teaching requires the acquisition and refinement of unique skills and knowledge. It brings a sense of dignity to teaching as people focus upon producing quality instructional materials, intellectually stimulating online discussions and relevant feedback on student assignments. Online faculty who embrace teaching as a craft will be more likely to operate by an internal standard of excellence that helps them to cultivate a work ethic and be a colleague who willingly shares best practices with others. Additionally, a growing sense of confidence is characteristic of reflective teachers because it enables instructors to avoid the paralyzing effect of always having to prove or compare themselves to others (Sennett, 2003).

Research studies on experts in distance education have found that skill development and developing expertise are tied closely to the timing, quality and quantity of deliberate practice. The use of mentors plays a vital role by providing guidance, monitoring progress and establishing appropriate goals that promote optimal growth. Bruning, et al (2004) noted that research indicates deliberate practice can help less talented people surpass the achievements of those who are more talented. Skill acquisition among young athletes, mathematicians and musicians indicates that individuals follow a similar learning process. The key is having the appropriate guidance and intentional practices that cultivate superior performance. "The best practice occurs under the watchful guidance of a skilled mentor who helps the developing expert set goals and monitor improvement" (Bruning et al, 2004, p. 177).

Professional staff development programs can utilize some of the principles found in effective mentoring through the peer review process. Online distance education faculty can benefit from the insights of a peer reviewer who can provide clarity to their work and being receptive to insights from the review on what the faculty member is doing well and what areas that for improvement is critical. Online teachers can be proactive by participating in conferences, reading literature, sharing with their colleague's best practices and instructional resources and sharing tips learned in peer reviews with others.

The University of Phoenix encourages its SAS online distance education faculty to refine facilitation skills and instructional practices using online peer reviews, and by encouraging faculty to participate in a variety of free online workshops that are held both online. The following list of workshops is a partial list of the professional activities that are available free of charge to online faculty members (Faculty Development Workshops, 2005):

> **New Student Facilitation: Helping Learners Succeed -** focuses on resources and strategies facilitators can use with new students. Addresses common characteristics of new SAS students and present methods for helping entry-level students develop core competencies.
>
> **Online Tone -** emphasizes importance of proper tone in an online classroom, focusing on the application of appropriate tone when responding to and offering feedback to students. It will enable faculty to identify and develop communication skills that are necessary for teaching in the online classroom.
>
> **Critical Thinking -** introduces the components of the critical thinking process and identifies various methods for teaching critical thinking skills. Topical areas include taxonomies and frameworks for understanding critical thinking, and cognitive abilities and affective dispositions in critical thinking.
>
> **Learning Teams -** facilitation techniques and evaluation methods. Participants will explore Learning Team processes, including conflict resolution, behavioral guidelines, and factors affecting team interaction.
>
> **Difficult Student -** various approaches for resolving several types of conflicts. Participants will examine team dynamics and strategies for helping students with team approaches.
>
> **Evaluating student writing -** develop clear writing assignments, assess student papers effectively, and help students improve their writing skills. The workshop provides a review of writing principles and includes materials to assist participants to establish clear grading criteria for written work and provide effective feedback to students.
>
> **Plagiarism -** provides the knowledge and tools necessary to detect plagiarism.
>
> **Student Evaluation -** addresses grading plans, criteria for grade changes and grade grievances, and qualities of effective feedback

Some online educators have legitimate concerns about peer review evaluations being accurate and objective toward the teaching and learning process. Educators fear having subjective reviews that are politically motivated or have excessive focus on minor administrative details (i.e. alternative email address in a syllabus). It is important that the review process be based on objective and measurable evaluation standards and that administered by trained personnel who understand the dynamics of the teaching. Teachers can profit from constructive insights into their online work and sharing with colleagues who relate to the challenges of distance education. Brookfield (1995) eloquently describes how reflective teachers maintain a sense of high academic expectations and a positive mindset toward the educational process of online successful acceptance and embracement of the results of an online peer review:

> *Critically reflective teachers learn from the past but live in the present with an eye to the future. Because they know that every class has its own dynamic, they cease to rely only on methods and activities that have worked in the past. Their practice is infused with a sense of excitement and purpose. There is a continual checking of assumptions, a continual viewing of practice through different lenses, and a continual rethinking of what works, and why. Knowing that each new group of students brings its own challenges, they see their life as lived in forward motion. Because tomorrow is unpredictable, there is always the chance for new learning from practice. (p. 265)*

## Implications and Conclusions

Online Peer Reviews can be a win-win situation for faculty and administration if the program is established in a similar manner to the rafting guide coaching the rafters. Fear of falling must be overcome by supportive administration, careful peer review selection and training, and communication to online faculty about the non-evaluative nature of the peer review. Online educators should be educated about coaching and helpful objectives of the online peer review. They should be encouraged to share tips and best practices with colleagues and be receptive to suggestions and changes. Administration should carefully plan and monitor any online peer review program to ensure that it is effective in meeting its goals. The successful online peer review is to help faculty fly instead of fall, and encourage online faculty to participate in faculty workshops for further development.

## References


Adams, J.R. (2002). *Implementing a collaborative support model for online faculty: a case study.* Dissertation. University of Virginia. Available at Proquest.

Anderson, J. R. (2005). *Cognitive psychology and its implications* (6th ed.). New York, NY: Worth Publishers.

Bloom, B.S., Bertram B. Mesia, & D. R. Krathwohl (1964). *Taxonomy of Educational Objectives* (two vols: The Affective Domain & The Cognitive Domain). New York. David McKay.

Brookfield, S. (1995). *Becoming a critically reflective teacher*. San Francisco, CA: Jossey-Bass.

Bruning, R. H., Schraw, G. J., Norby, M. N., & Ronning, R. R. (2004). *Cognitive psychology and instruction* (4th ed.). Upper Saddle River, NJ: Pearson.

Carr, J. F. (2005). Standards to practice. *Journal of Staff Development*, *26* (1), pg 48.

Eysenck, M. W. (2001). *Principles of cognitive psychology* (2nd ed.). New York, NY: Psychology Press.

Knowles, M. (1984). The Adult Learner: A Neglected Species (3rd Ed.). Houston, TX: Gulf Publishing. Retrieved March 3, 2005, from: http://tip.psychology.org/knowles.html

Peer Monitoring of TRIO Programs (2005). *Peer Monitoring of TRIO Programs Helps*

*Find Ways to Save Money; Ensures Compliance with Federal Regulations.*

*Northwest Association of Special Programs*. Retrieved, March 3, 2005, from: http://users.moscow.com/mareese/bestpractices/peer_review.html

Pushing for Peer-Review (2005). School Administrators Article. *Education World,* Retrieved March 3, 2005, from http://www.education world.com/a_admin/admin047.shtml

Russ, G. J. & Klug, M.G. (1999). Attitudes of business college faculty and administrators toward distance education: A national survey, *Distance Education*, *20* (1) 109-129.

Schacter, D. L. (1996). *Searching for memory: The brain, the mind, and the past*. New York, NY: Basic Books.

Sennett, R. (2003). *Respect in a world of inequality.* New York, NY: W. W. Norton.

University of Phoenix Online (2005). Faculty Development Workshops.

U.S. Department of Education, National Center for Education Statistics. (1999). Distance education at postsecondary education institutions: 1997-98. NCES 2000-13. Washington, DC: Author.

Wolf, W. J. (2003). Student Peer Reviews in an Upper-Division Mathematics Class, *Exchanges, The Online Journal of Teaching and Learning in the CSU*, September. Retrieved March 3, 2005, from http://www.exchangesjournal.org/classroom/1156_Wolfe.html

## About the Authors

**Kimberly Dr. Blum Ph.D.**

**Dr. Kimberly Blum** helped create the Online Peer Review Program in the School of Advanced Studies at the University of Phoenix, trains potential Online Peer Reviewers, and mentors online doctoral students and teaches doctoral research classes at the University of Phoenix, School of Advanced studies for the past four years, starting over six years ago in the IS&T Online department, as well as conducting quantitative and qualitative research studies. In addition, Dr. Blum trains online doctorate faculty telecommuting with virtual worldwide teams; previous experience includes international system installments and training. Dr. Blum holds a B.S. in Information Systems, Masters in Management of Organizational Management, and a Ph.D. in education researching for the final dissertation online distance education student communication patterns, learning styles, and barriers.

She may be reached via email at: kdblum@email.uophx.edu

**Brent Muirhead Ph.D.**

**Brent Muirhead** has a BA in social work, master's degrees in religious education, history, administration and e-learning and doctoral degrees in Education (D.Min. and Ph.D.).

Dr. Muirhead is the Lead Faculty and Area Chair for GBAM Business Communications in the graduate department at the University of Phoenix campus in Atlanta, Georgia. He teaches a diversity of undergraduate and graduate level courses in Atlanta and online. He is an Associate Editor for Educational Technology and Society and he has worked as a visiting research fellow to Robert Gordon University, Aberdeen, Scotland.

He may be reached via email at: bmuirhead@email.uophx.edu.

**Editor's Note**: Individual and group reflective thinking are the subject of this paper. Reflection supporting tools for computer-supported collaborative learning provide learning opportunities that are parallel to individual and collaborative activities in a classroom setting. Collaborative reflective thinking is a product of group sharing through discussion and other reflection supporting tools. Results show significant support for the reflective process, especially in computer-supported collaborative learning environments.

# Design and Analysis of Reflection-Supporting Tools in Computer-Supported Collaborative Learning

**Seung-hee Lee**

## Abstract

This study proposes design principles for reflection-supporting tools in computer-supported collaborative learning (CSCL) environments as well as ways to analyze how such tools influence group performance and the perception of learners on group learning. The functions of reflection-supporting tools suggested in the study were *group workplace*, *thinking sharing board*, *reflective journal*, and *reflective scaffolding*. When employed in group learning within CSCL environments, reflection-supporting tools turned out to positively foster learners' performance and process. In addition, learners felt that these tools made a significant contribution to meaningful reflection and inquiry within CSCL.

**Keywords:** Computer-supported learning, Collaboration, Reflection, Reflection- supporting tool

## Background

In computer-supported collaborative learning (CSCL), knowledge construction is produced by learners' active thinking and reflection. Higher-order thinking is essential for learning cycle in which learners can look back and monitor themselves on their own. The same is true when learners change their own cognitive structure during learning; when learners work in groups, they tend to modify and further expand both individual and group cognitions. Researchers such as Dewey (1933) and Schön (1983) call this type of thinking "reflection". Constructive learning theories and models in recent times emphasize the power of learners' reflective thinking.

Some of the CSCL studies (Dillenbourg, 1999; Koschmann, 1994; Stahl, 2002) indicate that meaningful group learning depends on facilitating thinking and idea sharing though peers' discourse. Throughout group reflection, learners in social contexts can experience negotiation and meaning elaboration. While previous studies have focused individual-oriented reflection, recent CSCL studies stress the roles of collaborative reflection for meaningful group learning.

The studies on reflection have suggested the need to design reflective inquiry fostering environments. However, these studies tend to belong to the categories of theoretical studies, or to focus on technical implementation of tools from the perspective of the developers, not those of the learners or instructors. Further, most studies of reflective inquiry fail to mention that instructional designers should design learning systems and support tools that can assist learners' reflective thinking in order to facilitate knowledge transfer (Felton & Kuhn, 2002; Häkkinen, Jarvela, & Dillenbourg, 2000; Lin, 2001). In fact, only a few studies (Bell & Davis, 2000; Kolodner & Guzdial, 2000) focused on such key issues.

Since collaborative reflection in group learning is regarded as an important component in CSCL environments, how to support collaborative reflection needs to be explored. Therefore, the purpose of this study is to suggest cognitive tools for supporting learner reflection in CSCL environments. Also, this study intends to empirically verify the effects of reflection-supporting tools on the group learning.

## Reflection in Learning

Meaningful learning results from balanced integration of experiences and reflective thinking. Learning activities without enough reflective thinking can bring superficial level of knowledge. In traditional classroom environments, learning sometimes occurs with the limited levels of reflection.

However, reflection plays a pivotal role in group learning where an effective activity is related to the individual cognition as well as the social interactions among learners (Kemmis, 1986). In CSCL environments, learners become naturally experienced not only with internal conflicts, but also with social conflicts from multiple perspectives of their peers. These kinds of socio-cognitive conflicts or psychological burdens occur frequently in CSCL environments, as compared to classroom learning. Reflection can resolve such conflicts and result in the equilibrium of learner's cognitive structure.

Also, reflection as a constituent in group learning process should be influenced by collaborative interactions. With collaborative reflection, learners can compare their own thinking with those of others, and if appropriate, adapt their abstract thinking towards different and perhaps more appropriate and meaningful learning goals.

The possibility of tool implementation for facilitating reflective thinking has only been recently studied to determine if such technologies or tools can support learning. For instance, some studies have introduced functions of self-monitoring questions (Kolodner & Guzdial, 2000) as well as knowledge representation tools to compare their opinions or solutions with peers by explaining what they understood (Bells & Davis, 2000). In addition, some researchers have provoked reflective thinking by using visualization methods (Kyza, Golan, Reiser, & Edelson, 2002), while still others have employed elaborate reflective devices (Balyor, Kitsantas, & Hu, 2001).

Repeatedly to say in the study, reflection by nature has a social aspect and is strongly influenced by the community activity. As many scholars have pointed out, social learning environments are significant for providing individual-level strategies (Lin, Hmelo, Kinzer, & Secules, 1999). Reflective discourse with peers and more experienced others can improve both self and group actions. Based on the review of previous studies of tool implementation, it seems that there is support for reflection in group learning, but how to support reflection in CSCL environments remains to be revealed.

## Design and Development of Reflection-Supporting Tools

Reflection support plays a role as a mediator for learners to construct collective knowledge. To foster depth within such knowledge construction, such cognitive tools should extend beyond traditional discussion board functions. The key design principles for reflection-supporting tools suggested in the study emerge from previous theoretical and developmental studies (Bell & Davis, 2000; Gama, 2000; Gutwin, 1997). The basic principles used for the study are: facilitating social awareness, thinking visualization, learner discourse, and group meta-cognition. Based on these principles, the major functions for reflection-supporting tools were identified and developed to promote collaborative reflection (see Figure 1). The major tool functions are described in the following sections.

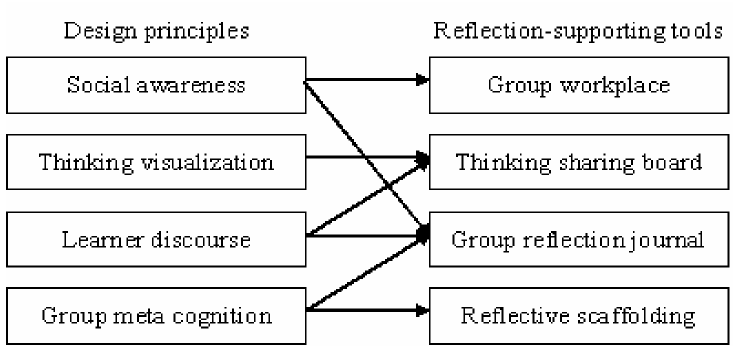



**Figure 1. Design framework of reflection-supporting tools**

***Group workplace*:** The group workplace is a transformed traditional online library, in which learners can categorize the mid-term or final products as well as accumulate project-related information and learning resources. Learners can save their work files in different places, according to different subject matters so that they can identify visually how the project products have developed over time.

***Reflection journal***: Learners can regularly record daily reviews on reflection journal pads while performing in their respective group projects. They can write what has been done so far, what should have been done better, and how much they have learned. With use of the group reflective journal, they can freely write their feelings, thought processes, and difficulties they experienced during the learning process. In addition, all records can be shared with other learners.

***Thinking sharing board***: This tool is intended to trigger learners' convergent thinking as well as divergent thinking. Tagging of ideas as pro, con, or neutral shows what stances learners can have on debate messages, and visualize how their own opinions and thinking in group work flows. Ideas, discussions, and debates of learners can be shared, compared, and contrasted among peer learners in-groups as well as between-groups.

***Reflective scaffolding*:** The purpose of reflective scaffolding is to stimulate learner's meta-cognitive activities related to problem-solving and for offering informative questions and cognitive assistances. Cues from this tool can be categorized into monitoring two parts such as subject domain and activity process. With reflective scaffolding, learners can look back at what they did while participating in projects. Group performance-related cues for reflection-supporting tools are provided during learning.

Those tools described above were developed to support collaborative reflection. With such tools, all learning products and processes are open and shared with peer learners at any time. For instance, learners can see or edit project products in the middle of working. They can also build their journals by collaboratively writing reflection notes within their journals.

## Impacts of Reflection-Supporting Tools

### 1) Data Collection and Analysis

The study was conducted with adult learners in three universities. For this study, 53, 48, and 50 participants respectively formed three groups and were distributed into groups which were provided with different reflection-supporting tools; (1) a group with collaborative reflection-supporting tools (CRG), (2) a group with individual reflection-supporting tools (IRG)[1], and (3) a group with no reflection support (NRG).

A pre-test was conducted to measure group differences and their homogeneity was confirmed. Data from the evaluations of group performance, surveys and semi-structured focus group interviews were analyzed using the constant comparative approach. After subjects experienced their online projects for five weeks, participants' accomplishments were evaluated by two content experts. They were also solicited for interviews related to their online experiences and their perceptions of the usefulness of various learning tools.

### 2) **Findings**

Within their group tasks, learners were asked to conduct group performance projects. To succeed, the participants of each group worked together to meet the common goal of completing the given projects. As shown in Table 1, the analysis of variance between groups (ANOVA) of the differences among the three groups' scores in the group performance was statistically significant. A post-hoc test results suggest that the CRG scored statistically higher than the IRG and the NRG.

**Table 1**

**ANOVA results for group performance**

| Group | N | Mean | SD | F | Significance |
|---|---|---|---|---|---|
| CRG | 53 | 62.66 | 3.87 | 19.51 | .00 |
| IRG | 48 | 54.60 | 5.64 | | |
| NRG | 50 | 51.83 | 3.66 | | |

Learner perception on performance was assessed with the reflection-supporting tools. Once again, the mean of score for group learning, CRG, was the highest when compared to the IRG and the NRG. Table 2 details the results of the ANOVA tests, which significantly favored the CRG on the awareness of the group performance.

From these results, we can find out that reflective thinking and reflection-supporting tools have a positive impact on group learning. In particular, the tools supporting collaborative reflection have a greater impact than those for individual reflection-support or no support regarding learner performance and perception.

Meanwhile, in order to understand how learners perceived the functions of reflection-supporting tools, the survey results of tool usefulness were analyzed. The findings show that both CRG and IRG were positive on their learning experiences and usefulness of reflection-supporting tools.

**Table 2**

**ANOVA results of perception on group performance**

| Group | N | Mean | SD | F | Significance |
|---|---|---|---|---|---|
| CRG | 53 | 48.37 | 6.74 | 2.56 | .00 |
| IRG | 48 | 44.04 | 7.16 | | |
| NRG | 50 | 45.16 | 7.27 | | |

At the beginning, learners with little online group learning opportunities seemed to have psychological and cognitive overload to collaborating online, but they perceived online learning positively as time gradually went by. According to the descriptive response from learners, they seemed to have unexpected meaningful experiences during the group projects. Learners from both GRG and IRG considered positively collaborative activities as learning experiences. The learner responses for user friendliness are noted below:

***Group workplace***: The CRG indicated that sharing 'group workplace' with other group members gave useful guidance and tips for conducting projects (92 percent), compared to the IRG (90 percent). 'Group workplace' seemed to help learners conveniently achieve their outcomes. At the same time, some learners were confused by the different boards per project subject. As noted in the quote below, a few learners seemed to feel cognitive overload, since they had to distinguish each of the functions, while conducting group projects.

> *I think work workplace have effective functions in terms of the fact that this tool gave learners independent boards per the different tasks or different subjects, compared to the traditional boards where all the files usually have to be upload in one board.*

***Thinking sharing board***: CRG mentioned 'thinking sharing board' was useful and positive (84 percent). They experienced deep reflection while group opinions were revised and well developed with the use of the embedded function. In particular, with tagging, they could relay the relevant messages under particular discussion topic categories, and that helped group discussions and negotiations become much effective. This is reflected in the following learner quote:

> *At the beginning, my group members showed different ideas and opinions so that we had a kind of conflicts in making consensus. But as time went by, we could naturally exchange multiple perspectives and accept others' ideas or revise them, even elaborate our own thinking better. I was quite surprised when I found my fellow's thinking and I felt the range of my thinking was much wide and reconsidered the project subject with better approaches. In particular, tagging to the messages on the board was a helpful function with that we could see how group ideas were developed and revised and finally consensus made.*

***Reflective scaffolding***: Regarding the preference for 'reflective scaffolding', learners responded that reflective scaffolding was helpful. For instance, one learner stated:

> *Reflective cues, questions or best practice for the group projects, seemed to play a good role for conducting projects. They were useful for us to guide group journal so that we were on the right tracks for learning.*

***Reflection journal:*** With 'reflection journal', learners became aware of the importance of reflective activities in the group learning projects. The learners in CRG perceived reflective activities important for group performance than learners of IRG (CRG: 76 percent, IRG: 68 percent).

On the other hand, the preference to individual reflection journal writing was slightly higher than collaborative journal writing. However, most of them agreed that it was of vital importance to have both their journal-writing available for critical reading and reviews by peers. As noted by one learner,

> *Without the reflective activities, we could hardly find out if our group projects have been on the right track, if we understood the essence of the projects or information, or if all the*

> *actions have been done as scheduled. I think that reflecting gave us the right judgment for good learning. Particularly with collaborative journal, we could monitor, find and revise what was missing or what was less good. We knew the fact that all the revisions are the required steps in the group learning and the focus of conducting projects was on the learning experiences and learning process.*

Interestingly, regarding reflection journal writing in particular, some learners in CRG answered that writing on reflection journals was not comfortable. It can be interpreted in two perspectives. First, they were not accustomed to put their records open to peer learners. These actions are more private in oriental cultural contexts. Therefore, even though learners realized the importance of sharing their reflective thinking, they still showed some feeling of hesitation to share what they had monitored in their online activities, until they got familiar with the benefits of the tool within online collaborative learning environments. Further studies have to address how to encourage learner internal motivation to open their thinking to others with deep consideration of their cultures and backgrounds.

Secondly, there was not full understanding on functions of the tools among learners, which led them unsure of how to apply them in writing journals. One of keys to effective learning performance and process would rely on not only providing useful reflection-supporting tools to learners, but also guiding them enough on how to use them in their group activities.

## Conclusion

Reflection is an in-depth practice in which learners participate in social behaviors such as communication or decision-making. They look back on their thinking process or actions, and listen to peer learners in groups or teams throughout their collaborative reflection. The social contextual environments with peer learners can be extremely useful in developing high-level cognitive structures. Previous studies have inclined to focus the area of theoretical reviews or technological implementation, but this study went further to draw out design principles for reflection-supporting tools and to identify the usefulness of several different learning and reflection tools. In this study, collaborative reflection-supporting tools turned out to positively impact group learning as expected.

Repeatedly to say, the findings show that reflection-supporting tools have positive impacts on the group performance as well as the perception of learners on collaborative learning. Also, the reflection-supporting tools in CSCL environments were effective and user friendly for group learning. This study clarified the ideas of Dillenbourg (1999) who argued that there is a need to externalize thinking and ideas with appropriate tools. Tools such as the project workplace, thinking sharing board with tagging, and reflection scaffolding that were proposed in the study were useful for facilitating reflective activities and for conducting group projects. On the basis of the results, follow-up studies on reflection support in CSCL environments are required with qualitative research approaches so that extra efforts should be put to identify changes in both learner behaviors and flow of thinking during group learning.

## References

Baylor, A., Kitsanstas, A., & Hu, H. (2001). Introducing the IPSRT and CPSRT(constructivist planning self-reflective tool): Tools to promote instructional planning for pre-service teachers. Paper presented at the Association for Educational Communication and Technology, Atlanta, GA.

Bell, P., & Davis, E. (2000). Designing Mildred: Scaffolding students' reflection and argumentation using a cognitive software guide. In B. Fishman & S. O'Connor-Divelbiss (Eds.), Fourth International Conference of the Learning Sciences (pp. 142-149). Mahwah, NJ: Erlbaum.

Dewey, J. (1933). How we think. Lexington: D.C. Health and Company.

Dillengbourg, P. (1999). Introduction: What do you mean by collaborative learning? In P. Dillengbourg (Ed.), Collaborative learning: Cognitive and computational approaches (pp. 1-19). Elsevier Science Ltd.

Felton, M., & Kuhn, D. (2002). Science as argument. In T. Koschmann, R. Hall, & N. Miyake (Eds.), CSCL 2: Carrying forward the conversation (pp. 493-498). London: Lawrence Erlbaum Associates, Publishers.

Gama, C. (2000). The role of metacognition in problem solving: Promoting reflection in interactive learning systems. Retrieved on May 21, 2004 from: http://www.cogs.susx.ac.uk/users/claudiag/edut ech/metacognition.html

Gutwin, C. (1997). Workplace awareness in real-time distributed groupware. Doctoral dissertation. The University of Calgary.

Häkkinen. P., Jarvela, S., & Dillenbourg, P. (2000). Group reflection tools for virtual expert community-REFLEX project. In B. Fishman & S. O'Connor-Divelbiss (Eds.), Fourth International Conference of the Learning Sciences (pp. 203-204). Mahwah, NJ: Erlbaum.

Kemmis, S. (1986). Action research and the politics of reflection. In D. Boud, R. Keogh, & D. Walker (Eds.), Reflection, turning experience into learning. NY: Nichols Pub.

Kolodner, J, & Guzdial, M. (2000). Theory and practice of case-based learning aids. In D. Jonassen & S. Land (Eds.), Theoretical foundations of learning environments (pp. 215-242). Lawrence Erlbaum Associates, Publishers.

Koschmann, T. (1994). Toward a theory of computer support for collaborative learning. Journal of Learning Sciences, 3(3), 219-225.

Kyza, E. Golan, R., Reiser, B., & Edelson, D. (2002). Reflective inquiry: Enabling group self-regulation in inquiry-based science using the Progress Portfolio tool. Proceeding for CSCL 2002 Conference. Retrieved on November 4, 2003 from: http://newmedia.colorado.edu/cscl/165.pdf

Lin, X., Hmelo, C., Kinzer, C., & Secules, T. (1999). Designing Technology to Support Reflection. Educational Technology Research & Development, 47(3), 43-62.

Lin, X. (2001). Designing metacognitive activities. Educational Technology Research & Development, 49(2), 23-40.

Schön, D. (1983). The reflective practitioner: How professionals think in action. NY: Basic Books.

Stahl, G. (2002). Contributions to a theoretical framework for CSCL. Proceedings of the Computer Support for Collaborative Learning (CSCL) 2002 Conference. Retrieved on November 5, 2003 from: http://newmedia.colorado.edu/cscl/81.pdf

## Acknowledgements:

This work was supported by
Korea Research Foundation Grant (KRF-2003-037-B00071)

## About the Author

**Dr. Seung-hee Lee** is a research associate at Kelley Direct Online Programs of Indiana University. She earned her doctorate from Hanyang University in Seoul, Korea in 2003 and has conducted post-doctoral research projects at Indiana University from September 2003-December 2004. Dr. Lee had worked previously in e-Learning Center of Korean National Open University in Seoul, where she had consulted and promoted faculty development for online teaching and learning. Her major research interests are online collaboration, reflective technologies, e-learning in higher education, and online mentoring.

Dr. Lee can be contacted at:

Kelley Direct Online Programs, 777 Indiana Ave. Suite 200, Indianapolis, IN 46202-3135 (Tel: 317-278-9084, E-mail: seuselee@indiana.edu)

### End Note

[1] Reflection-supporting tools for IRG feature the same functions but operate only for a particular learner. In such instances, learners might conduct group work but keep their private journals which are not shared with anyone.

**Editor's Note**: Each study adds to the storehouse of knowledge about distance learning. This data is of value to instructional designers, instructors, administrators, and other researchers. Some studies confirm what we already know – or thought we knew; some findings challenge our previous positions, others provide reinforcement.

Dr. Liu provides additional building blocks that support areas of significant difference and areas with no significant difference. The latter is as important as the former, because it gives assurance that little or nothing is lost if we implement learning programs online using appropriate pedagogy and technology. The flip side of the coin is ability to serve many learners who could not otherwise participate, and know that losses are not significant and significant gains can be achieved.

# Effects of Online Instruction vs. Traditional Instruction on Students' Learning

**Yuliang Liu**

## Abstract

This quasi-experimental study was designed to compare the effects of online vs. traditional instruction on students' learning in two different sections (online vs. traditional section) of a graduate course for K-12 school teachers on *Research Methods in Education* in the summer of 2003. The experimental group involved twenty-two graduate students who received online instruction on WebCT; the control group involved twenty-one students who received traditional instruction. Participants in both groups completed the same chapter quizzes and a final test, as well as other essay writings, peer critiques, and group projects during the 10-week summer semester. Results indicated the experimental group significantly outperformed the control group in most quizzes and the final test.

**Keywords**: online instruction, learning outcomes, significant difference, no significant difference.

## Introduction

Distance education has grown fast in recent years. In the 2000-2001 academic year, 56% of all 2-year and 4-year institutions in the United States offered distance education courses for various learners. In addition, 12% of all institutions planned to offer distance education courses in the next 3 years (Waits & Lewis, 2003). Currently, online instruction is a primary method for distance education. With online instruction, the student is separated from the teacher and connected through the use of a computer and the Internet. More and more institutions are offering online courses and/or programs to their students in order to meet various learners' needs. Online learning and instruction, as an integral part of the teaching and learning process in higher education, is growing as fast as the technology itself. On the other hand, traditional classroom instruction is face-to-face instruction, typically conducted in a classroom setting in a lecture/discussion/note taking mode.

Recent research has indicated that online education has positively influenced many aspects of education both directly and indirectly (CEO Forum, 2000; Phipps & Merisotis, 1999). Until recently, the viability of online learning was not proven. On one hand, Clark (1983, 1994) maintained that media do not influence learning in any condition. On the other hand, Kozma (1994) debated that educational technologies influence learning by interacting with an individual's cognitive and social processes in constructing knowledge. These earlier debates are still relevant since newly emerging technologies respond to the earlier criticisms and enable learners to use them more efficiently.

According to Phipps and Merisotis (1999) and Russell (1999), there have been two lines of research comparing students' end-of-semester grades or learning outcomes for online and traditional sections. The first line of research focused on the "significant phenomenon" and cited significant increases in learning outcomes for online learners over their traditional counterparts. The most widely cited literature in this line is McCollum's (1997) report. McCollum cited a sociology professor who divided his statistics class into two groups: one in online format and one in face-to-face (FtF) format. According to McCollum, online students had more collaboration and their performance outscored their traditional counterparts by an average of 20 percent.

Later studies also supported the "significant phenomenon". Day, Raven, and Newman (1998) compared and studied the effects of web-based vs. traditional instruction on students' achievement in undergraduate technical writing in an agricommunication course. They found that online students attained significantly higher achievement scores in the major class project and essay assignments than those in the traditional course. In addition, Day, Raven, and Newman found that online students obtained a higher mean gain in attitudes toward writing.

Nesler, Hanner, Melburg, and McGowan (2001) studied a large sample from 30 institutions and found that nursing students in distance programs had higher scores in professional socialization outcomes than their campus-based counterparts. Al-Jarf and Sado (2002) investigated two groups of freshman students in their first ESL writing course and found the experimental group (web-based instruction) made more gains in writing, became more efficient, made fewer errors, and communicated more easily and fluently, compared with the traditional classroom control group.

The second line of research supported the "no significant phenomenon." These studies cited no differences in learning outcomes between online and traditional groups. Navarro and Shoemaker (1999) found that about 90% online learners in a graduate MBA class believed that they learned as much as or more than they would have in a traditional classroom. Schulman and Sims (1999) did not find any significant differences on the posttest scores between the online and traditional students in an undergraduate course. Jones (1999) conducted a comparison study of an all web-based class to a traditional class and also found no significant differences in GPA between online and traditional learners.

More recently several other studies have found no differences in learning outcomes in various courses between online and traditional learner. Johnson, Aragon, Shaik, and Palma-Rivas (2000) compared a graduate online course with an equivalent course taught in a traditional format on outcome measures such as course grades and student self-assessment of their performance in the course. They found no significant differences between the online and traditional student groups. Less than significant, traditional students had slightly more positive perceptions about the instructor and overall course quality.

Ryan (2000) compared online and traditional student performance in construction, equipment and methods classes and found no significant differences in performance between the two groups. Student evaluations of the course were also similar. Similar results of no significant differences in performance were also found by Gagne and Shepherd (2001) in their graduate accounting class, as well as by Johnson (2002) in an introductory biology class.

Review of the above studies indicates most studies in this area found no significant differences in learning outcomes between online and traditional courses in various subjects. Fewer studies have been conducted at the graduate level.. This exploratory study was designed to investigate whether online instruction affects learners' learning during a semester-long graduate course in teacher education. Learners' progress in the online and traditional sections is assessed by chapter quizzes and final grades, as well as essay writings, peer critiques, and group projects. A pre-course assessment was conducted and analyzed to ensure that both sections were equivalent. Based on the above literature review, the major research hypothesis in this study was:

**Research hypothesis**:

> *There was no significant difference in learners' learning performance, as measured by chapter quizzes and final grades, between the online section and the traditional section, at the completion of a semester-long graduate course.*

# Method

## Participants

All students who self-selected to enroll in EDUC501 (Research methods in Education) in both online and traditional sections for 10 weeks in the summer semester of 2003 were solicited in the first week for participation in this study. EDUC501 is a required core course in education at the master's level at a midwestern state university in the United States. Students in this course were from different graduate programs in education. Twenty-four students enrolled in the online section, but two of them withdrew within the first two weeks due to time commitment and unexpected family issues. Thus twenty-two students in the online section were included for final analysis and twenty-one students enrolled in the traditional section. Thus, a totals of 43 participants in both sections were recruited to participate in the study. Participants in both sections were asked to complete consent forms and demographic surveys in the first week. A pretest of course content in both sections was administered. A preliminary analysis of the pretest revealed that the control group scored a little higher than the experimental group. No significant differences were detected in pretest performance between online and traditional sections.

## Instruments

Formative and summative assessments of participant learning were conducted in two major domains: knowledge and application. Knowledge assessment focused on individual learning and included seven chapter quizzes and one final test. Application assessment focused on collaborative learning and included a combination of essay writings, peer critiques, and a group research project. The application assessment is consistent with Wade's (1999) perspective. That is, writing is a unique indicator of student's learning including communication between student and student, as well as between student and teacher. The final grades of students were assigned based on these two major assessments. Both sections had the same quizzes, essay writings, and group research paper every week. Each chapter quiz was administered as an individual open-book test, but without peer discussion in both sections. Each quiz had 25 objective multiple-choice items regarding each chapter to be completed within 40 minutes. The quizzes in the online section were only available during a specific week and were graded instantly after the completion. Online learners were delighted to have immediate quiz results and feedback; quiz results and feedback in the traditional section were reported back to the class in the following week.

## Experimental Design

This study used a non-equivalent control group design. In both the experimental group (online via WebCT) and control group (traditional classroom), the dependent variables of learning performance were pretested and posttested. The independent variable was online vs. traditional instruction in a graduate course. Based on recommendations from the Institute for Higher Education Policy (2000) and Kearsley (2000), a hybrid of instructional techniques was employed in the online section. Specifically, several major features of WebCT were used throughout the semester such as weekly online writing, peer critiquing, bulletin board discussion, online testing, and e-mail. Constructivist learning theory was the major theoretical foundations for online instruction in this course. Instructional design was based on the ADDIE model (Analysis, Design, Development, Implementation, and Evaluation) proposed by Dick, Carey, and Carey (2001). For additional information on design, development, and instructional strategies used in this course, see other recent publications by the author (Liu, 2003a; 2003b).

To reduce learner anxiety and maximize learning, one FtF orientation was conducted in the first week for the online section. The traditional section met once a week for 3 hours and was primarily taught FtF throughout the semester. Both sections were taught simultaneously by the lead investigator in the summer semester of 2003. In order to make both sections as equivalent as possible, the instructional objectives, content, requirements, assignments, and assessments in both sections were the same.

### Procedure

The pretest was administered in paper-and-pencil format to both sections in the first week to determine initial learning and performance. Next, participants in the online section were introduced to the online WebCT environment from the second through the final week. Ongoing posttests, including chapter quizzes and final test, were administered online for the online section and administered in paper-and-pencil format for the traditional classroom.

## Results and Discussion

Pretests and posttests of learning performance in both sections were coded and analyzed using SPSS 12. Descriptive statistics of all quizzes and tests in online and traditional sections are presented in Table 1. Results of participants'seven chapter quizzes and one final test in both sections were analyzed using independent samples t test and presented in Table 2.

**Table 1**

**Descriptive Statistics of all Quizzes and Tests in Online and Traditional Sections**

| | GROUPS | N | Mean | Standard Deviation | Standard Error Mean |
|---|---|---|---|---|---|
| ch1 quiz | experimental group | 22 | 96.82 | 5.68 | 1.21 |
| | control group | 21 | 83.10 | 10.18 | 2.22 |
| ch2 quiz | experimental group | 22 | 92.73 | 4.81 | 1.03 |
| | control group | 21 | 88.33 | 8.56 | 1.87 |
| ch3 quiz | experimental group | 22 | 91.36 | 7.27 | 1.55 |
| | control group | 21 | 86.67 | 5.99 | 1.31 |
| ch4 quiz | experimental group | 22 | 90.23 | 7.48 | 1.59 |
| | control group | 21 | 83.10 | 7.33 | 1.60 |
| ch5 quiz | experimental group | 22 | 85.23 | 9.19 | 1.96 |
| | control group | 21 | 81.19 | 8.79 | 1.92 |
| ch6 quiz | experimental group | 22 | 89.77 | 8.66 | 1.85 |
| | control group | 21 | 86.90 | 4.60 | 1.01 |
| ch13 quiz | experimental group | 22 | 90.23 | 8.38 | 1.79 |
| | control group | 21 | 84.76 | 8.87 | 1.94 |
| Pre-assessment | experimental group | 22 | 41.4545 | 12.07 | 2.57 |
| | control group | 21 | 44.9524 | 8.66 | 1.89 |
| Final test | experimental group | 22 | 87.6364 | 7.24 | 1.54 |
| | control group | 21 | 77.7143 | 9.68 | 2.11 |
| Final grade | experimental group | 22 | 4.0000 | .00 | .00 |
| | control group | 21 | 3.8095 | .40 | .09 |

## Table 2

## Results of t Test in Various Assessments between the Experimental and Control Groups

| | | t-test for Equality of Means | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| | | | | | | | Lower | Upper |
| ch1 quiz | Equal variances assumed | 5.491 | 41 | .000 | 13.72 | 2.50 | 8.68 | 18.77 |
| | Equal variances not assumed | 5.423 | 31.033 | .000 | 13.72 | 2.53 | 8.56 | 18.88 |
| ch2 quiz | Equal variances assumed | 2.087 | 41 | .043 | 4.39 | 2.11 | .142 | 8.65 |
| | Equal variances not assumed | 2.061 | 31.178 | .048 | 4.39 | 2.13 | .05 | 8.74 |
| ch3 quiz | Equal variances assumed | 2.307 | 41 | .026 | 4.70 | 2.04 | .59 | 8.81 |
| | Equal variances not assumed | 2.318 | 40.159 | .026 | 4.70 | 2.03 | .60 | 8.79 |
| ch4 quiz | Equal variances assumed | 3.157 | 41 | .003 | 7.13 | 2.26 | 2.57 | 11.69 |
| | Equal variances not assumed | 3.159 | 40.969 | .003 | 7.13 | 2.26 | 2.57 | 11.69 |
| ch5 quiz | Equal variances assumed | 1.471 | 41 | .149 | 4.04 | 2.75 | -1.51 | 9.58 |
| | Equal variances not assumed | 1.472 | 41.000 | .149 | 4.04 | 2.74 | -1.50 | 9.57 |
| ch6 quiz | Equal variances assumed | 1.347 | 41 | .185 | 2.87 | 2.13 | -1.43 | 7.17 |
| | Equal variances not assumed | 1.365 | 32.307 | .182 | 2.87 | 2.10 | -1.41 | 7.15 |
| ch13 quiz | Equal variances assumed | 2.078 | 41 | .044 | 5.47 | 2.63 | .15 | 10.78 |
| | Equal variances not assumed | 2.075 | 40.555 | .044 | 5.47 | 2.63 | .14 | 10.79 |
| pre-assessment | Equal variances assumed | -1.087 | 41 | .283 | -3.4978 | 3.22 | -9.99 | 3.00 |
| | Equal variances not assumed | -1.096 | 38.129 | .280 | -3.4978 | 3.19 | -9.96 | 2.96 |
| Final test | Equal variances assumed | 3.818 | 41 | .000 | 9.9221 | 2.60 | 4.67 | 15.17 |
| | Equal variances not assumed | 3.792 | 37.013 | .001 | 9.9221 | 2.62 | 4.62 | 15.22 |
| Final Grade | Equal variances assumed | 2.222 | 41 | .032 | .1905 | .09 | .02 | .36 |
| | Equal variances not assumed | 2.169 | 20.000 | .042 | .1905 | .09 | .007 | .37 |

Results in Table 2 revealed that between online and traditional sections, no significant differences were found in chapter 5 ($\underline{t}_{(41)} = 1.47$, $p = .15$) or chapter 6 quizzes ($\underline{t}_{(41)} = 1.35$, $p = .18$). However, significant differences between both sections were found in all other five quizzes, including chapters 1, 2, 3, 4, 13, and the final test. Specifically, in chapter 1 quiz, $\underline{t}_{(41)} = 5.49$, $p = .00$; in chapter 2 quiz, $\underline{t}_{(41)} = 2.09$, $p = .04$; in chapter 3 quiz, $\underline{t}_{(41)} = 2.31$, $p = .03$; in chapter 4 quiz, $\underline{t}_{(41)} = 3.16$, $p = .00$; in chapter 13 quiz, $\underline{t}_{(41)} = 2.08$, $p = .04$; in the final test, $\underline{t}_{(41)} = 3.82$, $p = .00$. In terms of learners' final grades, $\underline{t}_{(41)} = 2.22$, $p = .03$. Thus, overall, the null research hypothesis described previously in this study was not supported.

In addition, regarding the students' perceptions and satisfactions of the course, the same students' evaluation form including 18 evaluation items was used by the lead investigator's department in both sections at the end of the course. Students' quantitative evaluation results revealed that the average in both sections was about the same (4.5 on a 5-point scale). However, student's qualitative comments indicate that students in the online section were more motivated than those in the traditional section. For instance, a few students in the traditional section complained about the content and frequency of chapter quizzes while those in the online section did not. In addition, students in the online section expressed greater satisfaction of the effectiveness of their learning in this course. A majority of students in the online section thought they had learned more in this course than from a traditional section. It was clear that such students' qualitative comments were consistent with the research findings described previously.

The results in this study indicate that there is a significant difference in learning outcomes between online and traditional learners. This study did not support the "non-significant phenomenon" described by Russell (1999). This finding is a surprise to the lead investigator due to various reasons. As described previously, the instructional requirements, activities, and content were attempted to be kept the same in both online and traditional sections. In addition, in the traditional section, the teacher also used various technologies such as using PowerPoint to present the course content in class and allowing students to access/print the teacher's chapter notes in Acrobat (.pdf) format from WebCT before the class. However, the results are consistent with the line of research called "significant phenomenon" described by Russell (1999). That is, this study supports prior research line called "significant phenomenon" in this area and indicates that online instruction can be a viable alternative for higher education since it can achieve better student learning or at least as well as the traditional instruction.

Results of this study are inconsistent with some prior research. This may be related to various reasons:

> First, a variety of samples were used. The samples in most such studies in this area were convenience samples. The sample in this study was also a convenience sample and participants were not randomly selected. Some studies involved undergraduate students while this study involved graduate students.
>
> Second, a variety of subjects were involved in such studies including accounting, nursing, and construction. In this study, a graduate educational research course was involved.
>
> Third, a variety of online instructional strategies were used. Some studies only used online writing assignments while this study used a combination of assessment techniques such as online quizzes/tests, writing, peer critiques, and group projects.
>
> Fourth, a variety of online technologies were used. Some studies used the normal course web site while other studies used specialized course management and delivery systems such as Blackboard and WebCT. This study primarily used WebCT for online course delivery. Care should be taken in generalizing results to other environments without further investigation..

## Conclusion

This study supports some previous research that (a) there is a significant difference in learning outcomes between online and traditional learners and (b) online instruction can be a viable alternative for higher education. This study has significant practical implications for higher education since many institutions are offering more online courses/programs. It also contributes to the current literature in the area of online instruction and e-learning. If online instruction is found to enhance student learning, more online courses/programs can be proposed. For example, embedded online courses may be used in place of more lengthy/costly traditional courses.

Due to various limitations of the study, care should be taken in generalization of results to other environments.

---

* An earlier version of this paper was presented at the International Congress of Psychology in Beijing, China, in August 2004.

**Acknowledgement: This project was partially sponsored by the Illinois Century Content Development Grant from Illinois Board of Higher Education from April 2002 to June 2003.

## References

Al-Jarf, A. & Sado, R. (2002). Effect of online learning on struggling ESL college writers. San Antonio, TX: National Educational Computing Conference Proceedings. (ERIC Document Reproduction Service No. ED 475 920).

CEO Forum (2000). The CEO forum: School technology and readiness report [Online]. DC: CEO Forum. Available: http://www.ceoforum.org/.

Clark, R.E. (1983). Reconsidering research on learning from media. Review of Educational Research, 53, 445-459.

Clark, R. E. (1994). Media will never influence learning. Educational Technology, Research and Development, 42(2), 21-29.

Day, T., Raven, M. R. & Newman, M. E. (1998). The effects of world wide web instruction and traditional instruction and learning styles on achievement and changes in student attitudes in a technical writing in an agricommunication course. Journal of Agricultural Education, 39(4), 65-75.

Dick, W., Carey, L., & Carey, J. O. (2001). The systematic design of instruction (5th Edition). New York: Addison-Wesley Educational Publishers, Inc.

Gagne, M. & Shepherd, M. (2001). Distance learning in accounting. T. H. E. Journal, 29(9), 58-62.

Johnson, M. (2002). Introductory biology online: Assessing outcomes of two student populations. Journal of College Science Teaching, 31(5), 312-317.

Johnson, S. D., Aragon, S. R., Shaik, N., & Palma-Rivas, N. (2000). Comparative analysis of learner satisfaction and learning outcomes in online and face-to-face learning environments. Journal of Interactive Learning Research, 11(1), 29-49.

Jones, E. (1999). A comparison of all web-based class to a traditional class. Texas, USA. (ERIC Document Reproduction Service ED 432 286).

Kearsley, G. (2000). Online education: Learning and teaching in no cyberspace. Belmont, CA: Wadsworth.

Kozma, R. B. (1994). Will media influence learning? Reframing the debate. Educational Technology Research and Development, 42, 7-19.

Liu, Y. (2003a). Improving online interactivity and learning: A constructivist approach. Academic Exchange Quarterly, 7(1), 174-178.

Liu, Y. (2003b). Taking educational research online: Developing an online educational research course. Journal of Interactive Instruction Development,

16(1), 12-20.

McCollum, K. (1997). A professor divides his class in two to test value of online instruction. Chronicle of Higher Education, 43, 23.

Navarro, P., & Shoemaker, J. (1999). The power of cyberlearning: An empirical test. Journal of Computing in Higher Education, 11(1), 33.

Nesler, M. S., Hanner, M. B., Melburg, V., & McGowan, S. (2001). Professional socialization of baccalaureate nursing students: Can students in distance nursing programs become socialized? Journal of Nursing Education, 40(7), 293-302.

Phipps R. & Merisotis J. (1999). What's the difference? A review of contemporary research on the effectiveness of distance learning in higher education. Washington, DC, USA: The Institute for Higher Education Policy.

Russell, T. L. (1999). The no significant difference phenomenon. Office of Instructional Telecommunications, North Carolina State University, USA.

Ryan, R. C. (2000). Student assessment comparison of lecture and online construction equipment and methods classes. T. H. E. Journal, 27(6), 78-83.

Schulman, A. H., & Sims, R. L. (1999). Learning in an online format vs. an in-class format: An experimental study. T. H. E. Journal, 26(11), 54-56.

The Institute for Higher Education Policy (2000). Quality on the line: Benchmarks for success in Internet-based distance education. Washington, DC, USA.

Wade, W. (1999). Assessment in distance learning: What do students know and how do we know that they know it? T.H.E. Journal, 27(3), 94-100.

Waits, T., & Lewis L. (2003). Distance education at degree-granting postsecondary institutions: 2000-2001. U.S. Department of Education. Washington, DC, USA: National Center for Education Statistics (NCES Pub 2003-017).

## About the Author

**Dr. Yuliang Liu** is assistant professor of instructional technology in the Department of Educational Leadership at Southern Illinois University Edwardsville. His major research interest is in distance education, online instruction, and research methodology.

Contact Data:

Yuliang Liu, Ph. D.
Department of Educational Leadership
Southern Illinois University Edwardsville
Edwardsville, Illinois 62026-1125 USA

Phone: (618) 650-3293     Fax: (618) 650-3808     E-mail: yliu@siue.edu