A Systematic Review of the Effectiveness and Efficiency of Networked ICT in Education

A State of the Field Report to the
Council of Ministers of Education, Canada and Industry Canada

Charles Ungerleider, Project Director
Tracey Burns, Lead Researcher

4 September 2003
Final edited version: 24 October 2003

The views expressed in this paper are entirely those of the authors.

Introduction

The Government of Canada has charged Industry Canada with a mandate "to help make Canadians more productive and competitive in the knowledge-based economy, thus improving the standard of living and quality of life in Canada" (http://www.ic.gc.ca/cmb/welcomeic.nsf/ICPages/Mandate). Connectedness — the ability to take advantage of information resources — is an important element in the Government's strategy:

> Connectedness is at the foundation of the knowledge economy and society. The speed and efficiency with which Canadians gain access to, and take advantage of, the Information Highway is of the utmost importance if we are to continue to foster a competitive Canadian presence in the global economy. Making sure that Canadians can access opportunities offered by the knowledge economy is also an essential factor in sustaining productivity growth and quality of life for all Canadians. (http://www.ic.gc.ca/cmb/welcomeic.nsf/532340a8523f33718525649d006b119d/f3785d91b380411f05256c6800556711!OpenDocument)

Industry Canada's approach to connectedness is based upon activities, programs, and policies related to the three pillars of the network age: infrastructure, use, and content. Connectedness is seen as a crucial part of an economy based on innovation by providing infrastructure and ensuring "that all Canadians have the means to participate in the creation and sharing of knowledge." It is the government's intention that "all Canadians must be able to access an affordable, world-class communications infrastructure in their regular day-to-day activities, as well as in times of emergency." (http://www.ic.gc.ca/cmb/welcomeic.nsf/532340a8523f33718525649d006b119d/f3785d91b380411f05256c6800556711!OpenDocument)

Lifelong education is a key ingredient in ensuring that Canadians are knowledgeable and productive citizens. In 2001, the Advisory Committee on On-Line Learning — a joint committee of the Council of Ministers of Education, Canada (CMEC) and Industry Canada — developed a research agenda related to on-line learning from which the topic for this review was derived.

1. Background

1.1 The research questions

In order to make informed and justifiable decisions about the allocation of scarce resources, Government needs research evidence about the use of information and communication technologies (ICTs) in lifelong education. Government has a broad range of criteria that it might use to evaluate the contribution of information and communication technologies to ensure that Canadians can take advantage of opportunities offered by the knowledge economy.

For instance, Government could assess information and communication technologies in terms of universality, effectiveness, efficiency, accountability, and choice. *Universality* is concerned with ensuring that all individuals have access to and can benefit from an educational experience. Universality obliges agencies to accommodate those whose location or physical, emotional, or intellectual characteristics might preclude their access to, or their achievement of the benefits from, the educational experience. To achieve universality, persons whose location or physical condition mitigates their access to educational opportunities may benefit from the use of information and communication technologies, just as ramps are used by persons confined to wheelchairs, material in Braille is used by those who are blind, and instructors who can sign are used by those who are hearing impaired. *Effectiveness* is concerned with the achievement of the intended objectives of the educational experience. *Efficiency* is concerned with producing the maximum benefits possible for the given expenditure of public monies. *Accountability* is concerned with reporting to the public about how resources it provided have been used to achieve the goals of the initiative. *Choice* is concerned with the latitude accorded the learner in decisions about the knowledge needed to realize his or her objectives. Although people might prize all of these values, all five cannot be fully realized simultaneously; they are incommensurable.

Given temporal and other resource limitations, Industry Canada in partnership with the Council of Ministers of Education, Canada sought answers to two categories of questions.

Category 1:

Is the research literature devoted to on-line and networked learning capable of answering questions requiring inter-subjectively testable hypotheses[1] according to the principles that guide the policy-oriented work of the Campbell Collaboration?[2]

- Is on-line and networked learning more *effective* than classroom-delivered instruction?
- Is on-line and networked learning more *efficient* than classroom-delivered instruction?

In answering these questions and to the extent that resources permit, we were asked to attend to the moderating influence of:

- **learner characteristics** such as: gender, age, language, prior education, physical and mental (dis)abilities; extent and nature of prior preparation for using on-line or networked resources.

---

[1] Using such hypotheses, the phenomena should be described in a sufficiently clear and detailed representation that others can reproduce the phenomena or observe the phenomena under the same conditions.

[2] Campbell and Stanley (1963) and Cook and Campbell (1979) identified factors that threaten the inferences that can be drawn from studies of interventions (or treatments) when conditions internal to studies make it difficult to support a claim of difference (internal validity) or that raise questions about the application of the findings to situations that differ from those initially studied (external validity). These factors include: (1) the impact of events occurring between administrations of the measurements used to assess change (history); (2) changes in the persons taking part in the study that would have occurred over time even in the absence of an intervention (maturation); (3) the influence of a prior test on the participant's performance on a second or subsequent test (practice); (4) changes in measurement procedures that might affect outcomes (instrumentation); (5) the tendency of scores to regress to the mean when groups of participants have been composed on the basis of extreme scores or characteristics (regression); (6) the bias introduced when participants volunteer or are selected for membership in a particular condition (selection); (7) the loss of participants while a study is being conducted; (8) selecting groups for comparison that may have changed or matured independently of the intervention (interaction of selection and maturation); (9) the tendency of persons being studied to perform better because they are being studied (John Henry or Hawthorne effect); (10) the fact that prior measurement might increase or decrease a participant's sensitivity to a subsequent measure (sensitization); (11) misattributing to other settings results that were produced because of the way the study was conducted (reactivity to experimental conditions); and (12) the difficulty isolating effects when participants take part in several treatments prior to their performance being assessed (interference).
Campbell, D.T. & Stanley, J.C. (1963) *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin Company.
Cook, T. & Campbell, D. T. (1979). *Quasi-Experimentation: Design and analysis issues for field testing*. Boston: Houghton Mifflin.
For Campbell Collaboration Guidelines for the preparation of Review Protocols (Version 1, January, 2001) see: http://www.campbellcollaboration.org/c2_protocol_guidelines%20doc.pdf

- **programmatic or situational characteristics** such as: duration of study (full- or part-time); system context (elementary, secondary, postsecondary, workplace training); geographic location (rural, urban); production features (text, audio, video, graphics, animation); granularity (modular, global), level of interactivity; nature and extent of the preparation of the instructional staff.

We were also asked a number of important subsidiary questions:

- What are the limitations of the literature devoted to on-line and networked learning as far as answering questions according to the requirements of inter-subjectivity and the principles that guide the policy-oriented work of the Campbell Collaboration?
- How might such limitations be overcome?
- Given the state of the field, what are promising directions for research on on-line and networked learning?

Category 2:

What implications can be drawn from the literature devoted to on-line and networked learning about the efficacy of government initiatives and partnerships?

Are there promising practices to which government should pay particular attention?

1.2 Definitions

What follows are our working definitions of the key concepts used in the systematic review:

On-line and networked technologies include all technologies that allow for interactive communication, such as e-mail, the Internet, networked systems, tele-teaching, and computer-mediated communication (CMC).

Learning was defined as the act, process, or experience of gaining knowledge or skill in any setting, including schools, the workplace, and intervention programs.

Effective was defined as producing or being capable of producing a desired effect — in this case, learning.

Efficient was defined as acting or producing effectively with a minimum of waste, expense, or unnecessary effort; also exhibiting a high ratio of output to input.

1.3 Other reviews

As this is an area of intense interest and activity, there have been a number of other related reviews. Most pertinent to our topic is the forthcoming meta-analysis by Bernard, Lou, and Abrami (2003), which compares empirical work on distance education and traditional education from 1985 to the present. This is a comprehensive research study that has been underway for over two years and as a result has an impressive number of citations. The studies reveal a small but significant positive effect of interactive distance education over traditional education for academic outcomes, and a negative effect for retention rate and student attitude toward subject matter. As this analysis focuses on distance education in its entirety and includes the research from 1985 onwards, many of the studies included in the Bernard, et al. review do not come under our purview, yet their analysis still raises many issues pertinent to our discussion.

Cavanaugh (2001) investigates the effectiveness of interactive distance education technologies in K–12 learning and conducts a meta-analysis of the available literature. She finds a small positive effect in favour of distance education on a number of dimensions, but does not specifically address achievement outcomes. Specifically, she finds a positive effect for small group size and programs of short duration, and a negative effect for primary instruction through distance.

Allen, Bourhis, Burrell, and Mabry (2002) compare student satisfaction with distance education to traditional classrooms in higher education and conduct a meta-analysis of available research. They find a student preference for traditional instruction, but no difference in satisfaction levels between the methods of instruction.

All of these studies offer interesting discussion points and will be returned to later. It is important to keep in mind, however, that these meta-analyzes focus on distance education as a whole, and as such include many methods that do not fall under the umbrella of networked education. In addition, they all focus on much broader time frames than our study (one reference is from 1947), and as a result are commenting on education that, even if it is networked, may be networked in a very different format than we would see today.

2. Identifying and Describing Studies: Methodology

2.1 Stage 1: Identifying studies

The best systematic reviews make a concerted effort to find all available work on the topic in question, particularly "grey" research (i.e., research that has not been published, most often due to null results). As a consequence, we have engaged in as thorough and innovative a search as we could devise, using both standard library search methods and directly reaching out to the international research community. The bibliographic details of all potentially relevant papers found through database, hand, Web site, and bibliography searches were entered into an EndNote database. The citations found in each search of a new source were checked against existing material and duplicates were excluded.

2.2 Reaching out to the international community

In order to find as much grey research as possible, we approached persons from 15 countries who are knowledgeable about information and communication technologies (see Appendix A for a partial listing of those contacted). Researchers from the following countries were contacted: Brazil, Canada, Chile, France, Guatemala, Hong Kong, Ireland, Japan, Mauritius, Norway, Paraguay, South Korea, Sweden, UK, and USA. Many of the researchers provided material for our review, either in the form of papers they had written (or a colleague had written), suggestions for contacts, or relevant Web sites. All suggestions were followed up and reviewed.

On May 5, 2003, we participated in a symposium about "Evidence-Informed Policy: Current and Future Research in E-Learning" convened in Ottawa by the Council of Ministers of Education, Canada and Industry Canada, the sponsors of this review. During our presentation of the plans for the review, we encouraged audience members — persons deeply knowledgeable about on-line and networked learning — to contact us with suggestions and citations. We followed up and reviewed the very few suggestions made.

2.3 Electronic databases

In addition to personally contacting researchers broadly knowledgeable about ICTs, our research team followed standard search protocol and conducted database searches for relevant work. As our goal was to have these searches be as multinational as possible, our research team was chosen for linguistic skill as well as research ability. Among our team we had members fluent in Mandarin Chinese, Russian, German, English, French, and

Italian. We also had access to Arabic, Spanish, Japanese, and Korean translators, if needed.

Team members searched 12 electronic databases: C2SPECTR (Campbell Conference search site), Cochrane Library, Dissertation Abstracts, EBSCOhost, Educational Index, ERIC, International ERIC (includes Australian Education Index, British Education Index), Canada Institute for Scientific and Technical Information Catalogue, PsychINFO, Sociological Abstracts, Social Sciences Citation Index, and StatsCanada Electronic Publications.

The search terms (keywords) used by the research team were: ICT and education, On-line learning, networked learning, on-line education, networked education, distance education, distance learning education, on-line instruction, elementary education, technology use and achievement, interactive learning environments, distance education and tele-learning, Web-based instruction, network-based instruction, technology in education, digital media in undergraduate education.

2.4 Hand, Web site, and bibliographic searches

The research team reviewed the contents of 85 journals (see Appendix B for the list of journals). The three members of the team were given a journal to check, and then reviewed their findings with the Head and Lead Investigator, given feedback on what their strategies were, and sent to look at the same journals again. After establishing that all three were using standardized criteria, they were all assigned the same three journals to search simultaneously: *Computers & Education*, *Journal of Computer Assisted Learning*, and *Computers in Human Behavior*. Analysis of their searches revealed that two members had a 92 per cent overlap with each other in their identification of relevant papers. The third member had only a 75 per cent overlap with the other two, due to being overly inclusive: he had found all the articles the other two had and more. As we were in the process of gathering studies to then be coded and weighed for relevance, we believed that his over-inclusiveness at this stage was not a problem. As a result, we felt confident that our research team was using a standardized search strategy and would uncover all articles relevant to the project. We then divided our list of 85 journals among the three research team members who conducted the hand search.

Bibliographies of papers identified through the above methods were also checked, and relevant papers obtained when possible.

2.5 Inclusion and exclusion criteria

To be included in the study, papers were required to have investigated the effectiveness and efficiency of networked and/or on-line learning as compared to standard classroom learning. By “investigated” we meant that the papers set out to test or compare (using any

methodology) the effects of networked and on-line learning versus traditional classroom learning. Papers that did not have an evaluative component such as —descriptions of programs, policy and discussion papers — were excluded. Given the speed at which changes in technologies occur and our limited time for the project, our research team was instructed to focus on the period from 2000 to 2003.

2.5.1 Screening 1

Screening was initially carried out on titles and abstracts alone. The research team screened all citations found through the electronic database searches. Random samples were also screened by other members of the research team to ensure reliability. Hand searches and bibliographic references, Web site, and personal contacts were also screened for inclusion in or exclusion from the review.

2.5.2 Coding

Full reports were obtained of all papers that passed the initial screening procedure. They were then coded on a number of different variables (see Appendix C for our coding sheet). Studies were classified according to the country where the study was conducted, how they were obtained, the type of intervention provided, characteristics of the participants (age, gender), and the methodology of the study (qualitative, quantitative). The presence or absence of certain other experimental criteria was also coded (for example, whether or not the study had a comparison or control group, whether the groups were assigned by random selection, and whether or not the person doing the rating was blind to the group to which the participant had been assigned). In addition, for quantitative studies, we coded whether or not an effect size was computed or could be computed based on the information provided.

2.5.3 Screening 2

After coding the full reports, papers were again screened for inclusion. Based on the information provided by the full report and using the same criteria as the first screening process, papers were again included in or excluded from the review. Papers that were included in the review based on the full paper were then entered into a database that recoded the coding categories for all of the variables of interest.

3. Identifying and Describing Studies: Results

3.1 Identification of reports

Table 1 numerically demonstrates our identification and coding process. A total of 5,894 potentially relevant articles were identified. Of these, 5,662 were excluded by screening abstracts and/or titles. Of the remaining 232 that were then read in full form and coded, a further 42 per cent (97) were then excluded, leaving a total of 135 that were included in the study.

Table 1: Identification of reports

	Number
Total number of possibly relevant articles identified	5,894
Number of studies excluded based on abstract or title	5,662
Screening 1: Met inclusion criteria on the basis of the abstract or title; full report obtained and coded	232
Screening 2: Coded full reports that met inclusion criteria and entered into mapping study database	135

4. Mapping the Relevant Studies

4.1 Characteristics of relevant studies

Papers deemed relevant to our questions were mapped using techniques similar to those of the Evidence for Policy and Practice Information and Coordinating (EPPI) Centre reviews. The EPPI-Centre is part of the Social Science Research Unit, Institute of Education, University of London (http://eppi.ioe.ac.uk/). Mapping requires the analysis of variables such as language of review (and country of origin), topic, population focus, study design, and specific keywords. A map provides a resource for a systematic description of research; a basis for any narrowing of inclusion criteria for a more detailed and quality assessed research synthesis; and a context for interpreting the results of the synthesis, including the nature of suggestions for primary research (Gough et al., 2003).

4.2 Countries in which the studies were conducted

Table 2 identifies the number and proportion of studies according to the country in which the studies were conducted. A majority (56%) were conducted in the United States. A much smaller percentage of studies (6%) were conducted in Canada. Some of the studies were done as collaborations between investigators in different countries and were thus counted twice, resulting in more than 100 per cent as a total. We have a total of 19 different countries represented by at least one paper each, which is due in part to the emphasis we placed on reaching out to international contacts.

Table 2. Country of study

	Number of Studies	Percentage of studies
USA	76	56
UK	12	9
Canada	8	6
Taiwan	7	5
South Korea	5	4
Australia	4	3
Israel	4	3
Singapore	3	2
Greece	3	2
China	3	2
Malaysia	2	2
Finland	2	2

Table 2. Country of study		
	Number of Studies	Percentage of studies
Germany	2	2
Ireland	1	<1
Japan	1	<1
Mexico	1	<1
New Zealand	1	<1
Sweden	1	<1
France	1	<1
Total Countries: 19		

4.3 Educational setting of the studies

Table 3 sets out the educational setting for the studies included in the mapping. As a study could be conducted in more than one setting, the total percentage is again more than 100. University level settings were by far (77%) the most common. A much smaller percentage of studies was conducted in secondary and elementary schools (9% each), and other locations such as the workplace, community centres, and street programs (another 7%).

Table 3. Educational setting of the studies		
Setting	**Number of studies**	**Percentage of studies**
University	104	77
Secondary School	12	9
Elementary School	12	9
Other (workplace, community centre, etc)	10	7
Total	**138**	**102**

4.4 Type of technology used in the interventions

Table 4 sets out the type of technology used in the interventions. As this review focused on on-line and networked learning, by definition all stand-alone systems were excluded. Within the realm of networking, however, there are still a number of different types of technologies. For descriptive purposes, we narrowed them into the following categories: Internet (some aspect of the intervention was conducted with Web-based materials); networked (such as intranet or networked computers within a school/workplace); tele-teaching; e-mail; computer-mediated communication; and other (includes all other systems, for example, “personal response systems,” a technology much like that used on

Who Wants to Be a Millionaire, wherein students are expected to "vote" on certain questions as the class progresses).

Table 4. Type of technology used in intervention		
Type	**Number**	**Percentage**
Internet	91	67
Networked	10	7
Tele-teaching	4	3
E-mail	17	13
Computer-mediated communication (CMC)	12	9
Other	13	10
Total	**147**	**109**

4.5 Type of methodology used by the experimenters

Table 5 provides a description of the type of methodology used by the experimenters. Studies were assigned to two categories: qualitative and quantitative. Within these categories, studies were further subdivided, with qualitative studies being broken down into case studies and surveys. The category of quantitative study was subdivided into randomized controlled trials (RCTs), quasi-experimental studies, and surveys. In addition, a third general category of "other" was created to encompass policy documents, reviews, and discussions pieces that did not involve an evaluative approach. As some interventions used more than one type of technology, the sum of per cents is more than 100.

Table 5. Type of methodology used		
Methodology	**Number of studies**	**Percentage of studies**
Qualitative	30	26
• Survey	11	
Quantitative	88	66
• Survey	22	
• RCTs (and other experimental manipulations)	59	
Other (discussion, policy, etc)	17	15
Total	**135**	**101**

4.6 Methodological features of studies

Tables 6 to 8 provide an overview of some of the methodological aspects of the studies. Because our review questions ask whether networked learning is more effective and efficient than traditional learning, any reasonable attempt to answer the questions should have, at minimum, a comparison group that allows for meaningful conclusions to be drawn regarding the role of the intervention. Additional experimental strengths include whether or not the assignment to experimental or comparison group was random, and whether or not the rater (of whatever outcome variable of interest) was blind to which participants were in the intervention group versus the traditional instruction group.

4.6.1 Studies using a control or comparison group

Table 6 outlines the number of studies included in the mapping that have a comparison or control group.

Table 6. Number of studies with control/comparison group

Control/Comparison Group	Number of studies	Percentage of studies
Yes	44	32
No	71	53
n/a	20	15
Total	**135**	**100**

4.6.2 Studies using random assignment to comparison or control groups

Table 7 demonstrates the number and percentage of studies with comparison groups that assigned participants to groups randomly. Random assignment, of course, allows for more confidence that any observed effect of the intervention is a function of the intervention and not, for example, a result of certain types of participants selecting the networked groups as a function of greater facility with and interest in computers.

Table 7. Number of studies (with comparison group) that used random assignment to groups.

Random assignment	Number of studies	Percentage of studies
Yes	7	16
No	37	84
Total:	**44**	**100**

4.6.3 Studies using a comparison or control group in which the rater was blind to the assignment of participants to groups

Table 8 illustrates the number and per cent of studies with comparison groups in which the rater was blind to the condition of the participant. That is, the number of studies in which the person rating performance (e.g., the test marker, the instructor) was not aware of which participants were in the experimental group versus the comparison group. This, of course, is a good way to control for experimenter bias (both pro and con).

Table 8. Number of studies (with comparison group) that had a blind rater.

Blind rater	Number of studies	Percentage of studies
Yes	2	5
No	41	93
n/a	1	2
Total:	**44**	**100**

5. In-Depth Review: Methodology

5.1 Moving from the mapping to the in-depth review

As part of our research plan, we intended to narrow the number of studies we coded for mapping purposes into a smaller group that would be better able to address our research questions:

- Is on-line and networked learning more *effective* than classroom-delivered instruction?
- Is on-line and networked learning more *efficient* than classroom-delivered instruction?

Both of these questions clearly require a comparison between on-line and networked learning and traditional classroom instruction. As a first criterion for inclusion in the in-depth review, we therefore required that the study offer a *meaningful* comparison between traditional classroom learning and networked and/or on-line learning. We stipulate *meaningful* in this case because merely investigating one of the two types of learning and then creating comparisons to the other on the basis of "well-known truths" or "things we know for sure" (without offering support for such claims) does not allow one to make a logically compelling argument. Our initial criterion for inclusion in the in-depth review was therefore that the study must have a comparison group to which learning could be compared.

As previously illustrated in Table 6, only 44 of the 135 studies in the mapping study met this criterion. We had initially planned to use a second criterion based on sound experimental procedure that would require the studies to have random assignment to the comparison groups. However, as only 11 of the 135 studies included in the mapping study and only 7 of the 44 studies that met our first criterion met this second criterion, it was decided to abandon any further criteria and instead include all 44 studies with a comparison group in the in-depth review. It is important to note here that we are not requiring random assignment or the use of a control group for inclusion in the in-depth review; rather, the presence of <u>any</u> comparison group (randomly assigned or not) is enough to merit inclusion. Closer inspection revealed that 7 of the 44 studies with a comparison group were actually pre-2000. As we had established a timeframe of 2000-2003, we eliminated these 7 studies, resulting in 37 studies included in the in-depth review.

If the data were in quantitative form suitable for statistical synthesis, they were considered for inclusion in a meta-analysis. Twelve studies were included in the meta-analysis. The synthesis for the remaining 25 studies took the form of a narrative summary or a qualitative review.

6. In-Depth Review: Results

6.1 Study characteristics

Of the 25 studies included in the narrative review, 19 (75%) originated in the USA. Australia and the UK each had two, and Taiwan, Finland, and Israel each had one (one study is a collaboration between Taiwan/USA, and is thus counted twice). Importantly, although one of the 8 studies from Canada included a comparison group and is included in the meta-analysis, none of the others qualified for the in-depth review. Twenty-four of the 25 studies were quantitative of some sort (either RCT, quasi-experimental, or survey). Of the 31 qualitative studies from our original sample, only one had a comparison group and thus met our criterion for inclusion in the in-depth review.

In terms of educational settings, 20 (83%) of the 25 studies in the narrative review were conducted in universities, 2 in secondary schools and 2 in elementary schools. Only 5 (21%) had random assignment to the groups. In addition, only two (8%) had a rater who was blind to the condition of the participants. Sample sizes ranged from 27 to over 500 participants, with two studies not reporting that information.

6.2 Table of studies

Author, Date, and Country	**Study Type**	**Aim**	**What was studied?**	**How?** Note: ES = Effect Size
Aldridge (2003) Australia	quantitative (survey)	Investigate learning environment and student outcomes in a grade 11 on-line Nuclear Physics course.	• sample: 32 grade 11 Physics students in two classes • perception of the learning environment in the two classes • knowledge of nuclear physics	Survey, pre- and post- tests Narrative review; t-tests (but no between comparison groups). ES: NO
Barile, Durso (2002) USA	quantitative	Determine whether using computers in group writing tasks is an effective mode of communication	• sample: 99 university Developmental Psychology students (volunteered for experiment) • communication at meetings to write a group term paper (either e-mail, 'net meeting' or face-to-face)	Grading, logs of communication Rater blind to condition! ANOVA on drafts ES: NO
Buchanan (2000) UK	quantitative	Investigate the efficacy of Web-based formative assessment	• sample: 214 university Psychology students (unequal N's: 16 vs. 200+) • use of study aid (PsychCAL) to non-use of aid (self-selected) • achievement on tests	Outcomes for the course Mann-Whitney U tests of means ES: NO
Clark, Jones (2001) USA	quantitative (survey)	Investigate the difference between traditional and on-line formats in a public speaking course	• sample: 61 university students in either on-line or traditional public speaking course (self-selected) • performance (self-report and independently assessed) in on-line class with performance in traditional class	pre- and post-test questionnaire; outcome assessment by 'expert' MANOVAs on survey data; MANOVAs and t-tests on independent assessment ES: NO

Author, Date, and Country	**Study Type**	**Aim**	**What was studied?**	**How?** Note: ES = Effect Size
Cooper (2001) USA	quantitative (survey)	Compare student characteristics in on-line and traditional classes	• sample: unknown number of students in two sections of the same university course • course evaluations for an academic year (course unclear)	course evaluations, survey descriptive statistics ES: NO
Dewhurst, Macleod, Norris (2000) UK	quantitative (survey)	Determine whether independent student learning can be accomplished with computers as well as traditional classrooms	• sample: 62 undergraduates in Human Physiology university course • attitudes to use and effectiveness of computer learning both pre- and post-computer-based instruction	Attitudinal survey (pre- and post-computer intervention) Descriptive statistics; Mann-Whitney U ES: NO
Dominguez, Riley (2001) USA	quantitative	Assess distance education courses by discipline for their effectiveness	• sample: many (number unclear) university students in both on-line and traditional courses • course results across the university by faculty to determine relative effectiveness by discipline	GPAs chi-square; Fischer's Exact test ES: NO
Dutton, Dutton, Perry (2002) USA	quantitative (survey)	Determine how on-line students differ from lecture students	• sample: 193 undergraduate engineering students in a computer programming course (on-line and lecture sections) • class performance and completion (retention)	Survey of attitudes, final course grade, completion rate Descriptive statistics, chi-square on survey data; regression co-efficients ES: calculable? (NO)

Author, Date, and Country	**Study Type**	**Aim**	**What was studied?**	**How?** Note: ES = Effect Size
Grabe, Sigler (2002) USA	quantitative	Evaluate the effectiveness of on-line studying	• sample: 191 university Introductory Psychology students (users and non-users of study tools) • performance on 3 tests	test scores ANOVAs ES: NO
Green, Gennteman (2001) USA	quantitative (survey)	Assess the impact of on-line courses	• sample: 57 university English composition students (in either a face-to-face or on-line class) • attitudes to Internet use and on-line course, performance in the course	attitudinal surveys, pre- and post-; GPA multiple significance tests (unclear what kind; only report p values) ES: NO
Hertz-Lazarowitz, Bar-Natan (2002) Israel	quantitative (survey, teacher evaluations, and outcome measures)	Determine effectiveness of Computer Mediated Communication (CMC) in elementary school writing	• sample: 599 grade 5 and 6 students in three learning environments: cooperative learning (CL), CMC, and CMC + CL • attitudes to writing; quality of writing portfolio	attitudinal survey, grades, written evaluations MANOVAs * gender, ethnicity analysis ES: NO
Johnson, Huff (2000) USA	quantitative (survey) qualitative	Look at effectiveness of CMC in distance education vs. traditional classroom	• sample: 76 Social Work students in two classes: Distance Ed and traditional • attitudes to computers and effectiveness of CMCs (e-mail and listservs)	Attitudinal surveys, open ended questions Narrative review; descriptive statistics ES: NO
Johnson, Aragon, Shaik, Palma-Rivas (2000) USA	quantitative (survey)	Compare learner satisfaction and achievement in face-to-face and on-line learning environments	• sample: 38 graduate students in either a traditional or on-line section of a HR development course • attitudes to on-line courses and support; performance outcomes	attitudinal survey; pre- and post-tests, final grades ANOVA, t-tests ES: NO

Author, Date, and Country	**Study Type**	**Aim**	**What was studied?**	**How?** Note: ES = Effect Size
Lockyer, Patterson, Harper (2001) Australia	quantitative and qualitative	Measuring effectiveness of Web-based teaching vs. traditional teaching in health education	• sample: 62 university health education students • observed learning outcomes, analysis of learner interactions, and perceptions of the experience	pre- and post- tests, in-depth interviews unclear what statistics ES:NO
Marttunen, Laurinen (2001) Finland	quantitative	The effectiveness of networked vs. face-to-face learning for argument styles	• sample: 46 university debating students (3 groups: face-to-face, e-mail, and a control) • improvement of argumentation skills as a function of course type	pre- and post-tests t-tests, ANOVAs ES: NO
Mason, Patry, Bernstein (2001) USA	quantitative (survey)	Examine the equivalence between computer-based and traditional testing	• sample: 27 undergraduate Introductory Psychology students • whether scores obtained from computer-based tests are equivalent to traditional methods	Attitudinal survey; test results ANOVAs of test results; descriptive statistics of survey ES: NO
Neuhauser (2002) USA	quantitative	Examine the effectiveness of on-line vs. face-to-face instruction as a function of learning style	• sample: 62 university students (not clear what course) in two classes: face-to-face vs. on-line • learning style and perceived effectiveness of mode of delivery	course grades, pre- and post-test surveys *Gender analysis descriptive statistics of survey; t-tests of grades, correlations ES: NO
Pérez-Prado, Thirunarayanan (2002) USA	qualitative	Compare student perspectives of on-line versus face-to-face section of a Teaching ESL course	• sample: 60 university pre-service teachers (two groups: on-line and traditional classes) • difficulties and benefits of Web-based instruction	interviews, journals coding for themes and ideas ES: NO

Author, Date, and Country	**Study Type**	**Aim**	**What was studied?**	**How?** Note: ES = Effect Size
Powell, Aeby, Carpenter-Aeby (2003) USA	quantitative (quasi-experimental)	Compare student outcomes focussing on disruptive students **(Does not provide a comparison group)**	• sample: 215 disruptive secondary students in two groups (computer based instruction (CBI) or teacher assisted CBI) • student outcome and achievement	GPA t-tests ES: NO
Sankaran, Bui (2001) USA	quantitative (survey)	Compare learning strategies and motivation on performance in Web-based instruction	• sample: 116 university business course (two groups: traditional and Web-based lectures) • student outcome and achievement, learning styles, and motivation in Web-based instruction	pre- and post- tests, survey of learning styles and motivation *ethnicity descriptive statistics, correlations, t-tests ES: NO
Schimmoeller (2003) USA	quantitative	Examine achievement and interest in science using a Personal Response System	• sample: 103 grade 5 science students (repeated measures comparison) • student achievement based on gender, type of question, and presence of intervention	achievement assessment, multiple choice *gender ANOVAs ES: NO
Sumner, Hostetler (2002) USA	quantitative	Compare computer conferencing and face-to-face communication in systems design	sample: 44 students in either a traditional meeting or e-mail conferencing assessing group scores on assignments, types of communication	confidence in group decisions; outcomes on assignments descriptive statistics ES: NO
Thirunarayanan, Pérez-Prado (2002) USA	quantitative	Compare Web-based and classroom-based learning in ESL teaching	• sample: 60 university preservice teachers • achievement based on format of instruction	pre- and post- tests of achievement t-tests ES: NO

Author, Date, and Country	**Study Type**	**Aim**	**What was studied?**	**How?** Note: ES = Effect Size
Wolverton, Wolverton (2003) USA	quantitative (survey)	Examine the effectiveness of on-line instruction vs. traditional course delivery	• sample: 65 university financial students (comparing experience with previous face-to-face experiences) • Survey data based on perceptions of the learning experience, instructor, and effectiveness	post-course surveys asking for comparisons with previously taken Face-to-face courses descriptive statistics ES: NO
Yu, Yu (2001) Taiwan, USA	quantitative (survey)	Examine impact of e-mail on the learning process, achievement, and attitudes	• sample: 68 university preservice teachers in a Computer and Education class (two groups: e-mail and non-e-mail) • Achievement and attitudes of the use of e-mail as a supplementary aid to instruction	post-test achievement indicators and attitudinal surveys ANOVAs, descriptive statistics ES: NO

6.3 In-depth review: Findings and methodological quality of studies

The 25 studies that met our inclusion criteria but did not include enough statistical information for a meta-analysis are summarized in the following in-depth narrative review. We attempted, by contacting the authors, to obtain the statistical information for studies that appeared to have appropriate statistical analyses for inclusion in a meta-analysis but did not report that information. Unfortunately, none of the authors responded to our request for further information.

Themes from studies that addressed similar questions were brought together and summarized in the discussion section following the narrative review. Within the review, studies were given a weight (high/medium/low) with which to assess their contribution to the discussion. The weighting system was based on that used by the EPPI Centre, which outlines three dimensions and one overall weight (Andrews et al., 2002):

A. Soundness of methodology
B. Appropriateness of study type to answer the review question
C. Relevance of the topic focus to the study of the review question
D. Overall weight of evidence that can be given to the study's results

Studies were rated as either high/medium/low for each of variables A, B, and C, and then combined to produce D, the overall weight given to the evidence. These weights were taken into account when summarizing data, combining themes, and identifying gaps in the research as part of the narrative review.

6.3.1 Narrative review: Summary of the studies included

Aldridge (2003) studied 32 grade 11 students and their perceptions of an on-line Nuclear Physics course. Students were given pre- and post-tests assessing their knowledge of the subject, and were also given an attitudinal survey regarding their perceptions of the learning environment. This study suffers from some important methodological flaws. Although students were given pre- and post-tests, they were not taught any nuclear physics prior to the course. Thus, the test showed only whether they learned anything from the course, not whether learning was commensurate with other methods. Achievement was not compared to any other learning paradigm. In terms of the attitudinal survey, students were asked to compare their experiences with those of traditional classrooms, but they were not asked particularly pertinent questions. For example, students said that they used a computer more often in their on-line course than in traditional courses — something one might reasonably expect from such an intervention. Similarly, although the experimenters were careful to gather information and place it in a context of learning and learning styles, they generally seemed to be missing the point. The narrative review of the study revealed many of the same themes highlighted in other studies: Students appreciated the use of computers and enjoyed group work and interaction in the networked environment. Some students, however, did not appreciate the computerized environment and felt uncomfortable learning in the on-line environment, causing the authors to caution that this type of instruction is not well suited

to some learning styles (although many of the other papers provided evidence that students of diverse learning styles perform and achieve well in networked environments).

Barile and Durso (2002) studied 99 university Developmental Psychology students' communication during meetings. Students were assigned a group term paper in one of three groups: Face-to-face, e-mail, or synchronous computer-mediated communication (CMC). Students were randomly assigned to one of the three groups, which helped reduce the possibility of the influence of personal preference for communication styles and also the influence of prior friendships. In addition to random assignment, this is one of the few papers in which the instructor grading the assignments was blind to the condition to which the students had been assigned, reducing the possibility of rater bias. Papers were rated on outcome (grades for the paper and a rough draft), and the interactions of the group were also monitored for communication styles and interactivity. Overall, no significant difference was found between the three groups, although the e-mail group did have a significantly lower score than the face-to-face group on the rough draft (the CMC group fell somewhere in between). In addition, the e-mail group spent more time discussing coordination of meetings and work, suggesting that they were having more trouble coordinating than the other two groups. The e-mail group also failed to attend to questions asked by the other group members. The authors concluded that given their results, e-mail is not an effective way to write collaboratively. However, no significant differences were found between the face-to-face and CMC groups, suggesting that collaborative writing with either method results in a similar process. This is an example of a good empirical study in this domain: a reasonable sample size, comparison groups, random assignment, blind rater, and appropriate statistical analyses. Despite this, the authors did not include enough statistical information (means, standard deviations) and so we were unable to include this study in our meta-analysis.

Buchanan (2000) studied the use of on-line study tools in an Introductory Psychology class. In all, 214 university students were given the option of using a set of on-line study tools to prepare for their exams. Because use of the tools was voluntary, there were dramatically different numbers of students in each condition: 16 students used the study tools and 198 did not. Not surprisingly, they found that the students who used the study aids scored significantly higher on at least some (but not all) measures of learning. Unfortunately, as membership in these groups was completely voluntary, it is not hard to imagine a scenario where the best and most motivated students used all available means to study for the exam, and thus were labelled "Users" and performed better on the exam — <u>not</u> because of the study tools themselves, but because these were the best and brightest students. (See comments on Grabe et al. below for a way to avoid this threat to validity.) In the absence of other solid data, strong arguments about the effectiveness of the on-line study tools cannot be made.

Clark and Jones (2001) studied 61 university students in either an on-line or traditional section of a public speaking course. They compared performance (both self-assessed and independently assessed) across the two conditions by means of a pre- and post-test questionnaire and independent assessment by an expert They found that both self-reports and the expert reports showed no significant difference between learning public speaking

in either of the two formats. Students reported enjoying the on-line format for flexibility in time and locale. Students in the traditional format reported enjoying the interaction with their classmates and the feeling that they learned as a group (seen as important for a subject such as public speaking). The authors pointed out that, although learning seems to be equivalent across the sections, membership in the groups was not randomly assigned. Thus, although these results demonstrated that students can learn equally well in both sections, the authors caution that it may not be true of all students in general; that is, students may have selected the format that most suited their learning style. In particular, they noted that on-line sections were chosen more often by males, who did not perceive group learning as important as females did, and by students who had access to a computer. In addition, the attrition rate was substantially higher in the on-line sections than in the traditional sections, a finding echoed by much of the other research presented in this review.

Cooper (2001) looked at course evaluations for a computer applications university course given in two sections: on-line and traditional. She does not include the number of students answering the evaluations, but does provide a breakdown of student characteristics: students in the on-line classes were in general older (but not substantially) than in the traditional class. Equal proportions of males and females were observed in both classes. The biggest difference between the classes was that students in the on-line classes were far more likely than students in the traditional classes to work full-time (56% of responders compared to 33%). Not surprisingly, students in the on-line class reported being happier with the flexibility of time in scheduling their classes. Also, they commented that the on-line section allowed them to lower indirect costs of education (e.g., travel, babysitters). However, more students in the traditional section reported being satisfied with the course. On the whole, this study offers too little in terms of methodological detail and discussion to be particularly useful. The fact that the author did not report an N anywhere in the report is cause for concern, whether due to negligence on the author's part or length restrictions on the part of the publisher. Although this study is very relevant to our topic question, it was given an overall weak weight due to confounding variables.

Dewhurst, Macleod, and Norris (2000) examined 62 undergraduate Human Physiology students' attitudes about computer learning. Pre- and post-computer intervention questionnaires were administered to the class, and perceptions of usefulness and effectiveness of computer-based learning (CBL) were assessed. The survey found a substantial shift in attitude post-intervention, reflecting more positive student attitudes to CBL. The authors noted that on the pre-test students rated themselves as either comfortable with computers or rather afraid of them. By the post-test, almost all of the students rated themselves as comfortable with computers, although a small minority still reported what they call "technophobia." The authors remarked that it is encouraging to see a positive change in attitudes through experience with computers, but that there may be a small group of students disinclined towards using CBL. Overall however, students rated their learning experience as enjoyable and believed that they learned as much by this method as they did in traditional lecture classes. Despite the positive attitudes expressed by the students, the authors were cautious in recommending changes to

curriculum. Students reported that they enjoyed the class but did not see it as an improvement over traditional class methods, and the majority of students still believed they would prefer at least some instruction in traditional lecture form. In addition, the students reported enjoying the flexibility CBL afforded them, but also said that they liked the structure a traditional lecture gave to their day. The authors argue that there is a role for structured and limited use of CBL to augment traditional classrooms, but that there is probably opposition to changing toward a complete CBL curriculum.

Dominguez and Riley (2001) studied student performance on a semester-by-semester basis in both on-line and traditional format courses. According to the authors, most of their university students take both formats during their time at the university. By looking at GPAs in both lower level on-line courses and higher level courses, they argue that the on-line courses prepare the students as well as the traditional lectures, with the exception of courses in one faculty: Management. Although they speculate why this faculty alone seems to have poor on-line performance compared to the rest of the faculties, they do not offer any substantive arguments to elucidate the pattern of results. They also do not offer means and numbers of students in both sections. This study appears to be an example of work that has appropriate statistical analyses and design (with the exception of a questionable call to probe a result of marginal significance), but poor interpretation and discussion, with the result that it does not offer very much to our discussion.

Dutton, Dutton, and Perry (2002) looked at characteristics of on-line students versus traditional students. Unlike other studies that show no difference between student characteristics, they found that students in the on-line class were significantly more likely to be older and significantly less likely to be enrolled in a traditional undergraduate degree program. On-line students were significantly more likely to work during the semester, live farther away from campus, and have previous computer expertise. In line with other studies, on-line students reported less need to see the instructor face to face, and a greater desire to control the flexibility and timing of their lectures. The authors also looked at course performance and found no significant difference between traditional students and on-line students in their programming course, which is surprising given that the on-line students rated themselves as more proficient with computers prior to the class. Of note are two findings relating to the environment of study: working had a significant negative effect on performance, although more so for traditional students than on-line students. In addition, on-line status decreased the probability of completion by 20 per cent.

Grabe and Sigler (2002) studied the use of on-line study tools in an Introductory Psychology class. They gave 191 university students the option of using a set of on-line study tools to prepare for each of three exams. The experimenters tallied the number of users (60–56%) versus non-users (40–44%) throughout the semester and compared their performance on exams. As the term progressed, fewer students accessed the on-line study tools and those that did spent less time using them. No statistics are offered regarding whether or not these were significant patterns. Similarly, no statistics are offered to determine whether the same students access the tools repeatedly or not (as students were required to sign in with an individual code, this information was available to the

experimenters). Overall, they reported a significant difference in performance according to the use/non-use of the tools, with users scoring higher than non-users. Unfortunately, as membership in these groups was completely voluntary, it is not hard to imagine a scenario where the best and most motivated students used all available means to study for the exam, and thus are labelled "Users" and perform better on the exam — not because of the study tools themselves, but because these were the best and brightest students. A much stronger argument would be made if use/non-use of the study tools could be connected to overall grade point average (GPA) and performance in other classes. If, for example, students with lower GPAs in other classes used these study tools and scored higher than their average GPA, it would be more reasonable to attribute this success to the study tools. In the absence of that data, however, we cannot make any strong arguments about the effectiveness of the on-line study tools, and this study receives a low weight for our review.

Green and Gennteman (2001) studied 57 university students in an advanced English composition course. They administered pre- and post- surveys and looked at GPAs of students in either an on-line or a traditional format course. Students enrolled in the sections did not differ in age or gender. The authors report only the survey responses of the students in the on-line section, leaving us unable to comment on the comparison between the two groups. The authors do report, however, that there was no difference found in the GPAs or retention rate as a function of format. This study is a good example of a relatively strong study design that does not live up to its goal: by reporting only the survey responses of the on-line group, we lack the comparative information needed to assess relative effectiveness.

Hertz-Lazarowitz and Bar-Natan (2002) investigated the effectiveness of computer-mediated communication (CMC) in elementary school writing. The 599students in grades 5 and 6 were studied in three different learning environments: cooperative learning (CL), CMC, and the combination of CMC and CL. The students reported on their perception of, and attitudes toward, writing. Their teachers evaluated their students as writers, and scored their portfolio of writing outcomes. Students rated themselves. Pre-tests showed no significant differences between the groups of students, while post-tests showed significantly better performance by the CMC+CL group (compared to the other two groups, which did not differ from each other). The experimenters investigated the role of ethnicity and gender, and found that there was no significant difference overall in performance by Arab and Jewish students (self-rating). Similarly, there were no significant differences overall for the performance of boys and girls (self-rated). In terms of teacher ratings, both computer groups (CMC and CMC+CL) were rated equally, and higher than writing without a computer (CL). Overall, the researchers comment that despite a lack of tradition in the use of computers in Language Arts teaching, this study gives evidence to support their usefulness.

Johnson and Huff (2000) studied 76 Social Work students enrolled in either a distance education class or a traditional lecture format. They were concerned about the use of CMC (in this case, both e-mail and listserv) and students' perceptions about computers in the two formats. Not surprisingly, they found that students in the distance education class

were significantly more likely to use CMC than students in the lecture class, and were also more likely to rate themselves (at the end of term) as more comfortable using the technology. No outcome measures or perceptions of achievement were investigated. Some of the comments echo those of other studies: students' main complaints about the distance education experience were difficulty accessing the technology and not being comfortable with the medium. However, students did appreciate being able to interact with other classmates via the listserv.

Johnson, Aragon, Shaik, and Palma-Rivas (2000) assessed comparative data obtained from students enrolled in one of two versions (face-to-face and on-line) of a graduate HR development course. Students were assessed prior to the start of the course and demographic data compared: no significant differences were found in age, year of graduation, work experience, or achievement. Students were then surveyed (as the course progressed) on their perceptions of satisfaction, learning outcomes, and self-assessment of ability. Students in the face-to-face group scored significantly higher in measures of satisfaction with their course and satisfaction with peer interactions. The face-to-face group also scored significantly higher in most other attitudinal measures, such as student-instructor satisfaction, instructor support, and departmental support. However, the quality of projects produced did not differ significantly across groups, nor did their grade distribution. Overall, this is a strong study, well-designed, appropriately analyzed, a relevant discussion. The authors seem to try a little too hard to dismiss the differences found between the groups in terms of satisfaction, but do make two clear points: 1) effectiveness of face-to-face versus on-line teaching (as measured by scores) does not differ across the two groups; and 2) lowered student satisfaction in on-line courses with support from instructor and department suggests that future on-line offerings be modified to ameliorate the situation. In particular, they suggest that instructor feedback on student progress be modified to be more rewarding for the student.

Lockyer, Patterson, and Harper (2001) studied 62 university health education students in both a traditional and Web-based classroom. Students were assigned to one of the two formats initially, and then switched to the other format halfway through the term. Students were first given a pre-test on a health-related subject (e.g., HIV/AIDS, nutrition), then given a tutorial on that subject in either a Web-based or face-to-face format. They were then given a post-test which tested knowledge for that subject area. The authors report the results of their statistical analysis in an unusual and incomplete manner, which makes them difficult to understand. In effect, there was no statistically significant difference between the groups — in one topic (HIV/AIDS), the Web-based students showed significant improvement in knowledge while the face-to-face students did not; and in the other topic (nutrition), the reverse was true. The authors also provide qualitative data regarding the students' experiences, which also suffer from the same limitations as the quantitative data: the data are presented in a confusing manner. Specifically, the discussion centres on Web-based learning (without comparison), so while we are told that the students found the Web-based section effective and interesting, we are not told if this is also true for the face-to-face format. However, some comments resonate: students in the Web-based section appreciated the flexibility, the time for reflection, and the opportunity to retain a copy of the interaction provided by that mode

of instruction. However, they also reported missing the opportunity to ask questions and interact with the instructor.

Martunnen and Laurinen (2001) looked at the effectiveness of face-to-face versus networked learning for a university debating class. Using an experimental design (although without enough statistical information for inclusion in the meta-analysis), students were assigned to one of three groups: a control, which did not receive any training in argumentation skills, and two experimental groups that did (either face-to-face or e-mail). Students were tested on their argumentation skills with pre- and post- tests, and a random sample of students drawn for an in-depth interview. The results demonstrated that students in the e-mail group learned to identify and choose relevant components of an argument, while those in face-to-face learning learned to put forward counter-arguments. Students in the control group learned none of these skills. The authors suggest that debating skills can be improved by both networked and face-to-face learning, but that the different environments develop different skills.

Mason, Patry, and Bernstein (2001) studied 27 undergraduate Introductory Psychology students' performance on traditional versus computer tests. The authors pointed to the concern that computerized tests are seen more negatively than paper-and-pencil tests, and address this in a tightly designed study. Over the course of a semester, students were given 10 tests, 5 paper and pencil tests and 5 computer-based ones (counterbalanced across the class). No significant differences were found in mean scores of paper and pencil tests compared to computer-based tests, leading the researchers to conclude that computerized testing is a valid and efficient way to test performance. They do note, however, that their computerized tests addressed the concerns most generally raised about on-line tests: the tests were programmed to allow changing answers, returning to previous questions, skipping questions temporarily, and so on. Thus, they caution that confidence in computer-based tests is merited as long as the computer-based tests allow the user the same flexibility and control as pencil-and-paper tests.

Neuhauser (2002) investigated the performance and attitudes of 62 university students enrolled in either an on-line or traditional lecture-based section. The author did not identify what course was being taught and, in fact, had a confound in her comparison group: the group identified as face-to-face received regular e-mails containing review materials. More importantly, the face-to-face group was required to take all three exams through e-mail. It is not clear that the face-to-face group was, in fact, a suitable comparison for the Web-based group. This flaw continues throughout the analyses, as the only significant differences that were found were related to the perceived effectiveness of the test and reviews, with on-line students rating them as significantly more effective than face-to-face students and performing better on tests (although at a non-significant level). One interesting aspect that the authors attempted to address was the performance of introverted students in a class that required active participation. Many researchers have suggested that on-line courses allow introverted students to feel more comfortable participating in class forums. In this study, however, the authors found no significant differences between student performance on-line and face-to-face on any of their measures of introversion and extroversion.

Pérez-Prado and Thirunarayanan (2002) conducted a *qualitative* investigation of students' views about on-line and traditional formats of a course (see the description of the *quantitative* study by the same authors cited in reverse order below). Comparing 60 pre-service teachers in an ESL instruction course, they provided journal entries and interviews to portray students' perceptions of the difficulties and benefits of Web-based instruction. One student in the off-line section spoke about her frustrations with using technology and noted that she learned better when familiar with the technology. Students in the on-line section offered both positive and negative comments: common positive themes include the feeling of being "prepared for the future," feeling more able to express themselves in the on-line environment, and enjoying the flexibility of learning on their own schedule. Negative themes included feeling at a loss without class structure, and the difficulty involved in group work if members do not check their e-mail regularly. Another aspect of the analysis was particular to the content being studied. As teachers being prepared to teach English as a Second Language, both the students and their instructor were concerned about a lack of emotional involvement in the on-line section of the course. The authors concluded that courses that require students to develop empathy or affective orientations may not be suitable candidates for Web-based education.

Powell, Aeby, and Carpenter-Aeby (2003) compared student outcomes for 215 disruptive secondary students as a function of the presence or absence of teacher-facilitated computer-based instruction. The study does not meet our criterion for an in-depth review in that it does not compare performance with non-computer-based instruction. However, because it focused on a special population of students, we included it in our narrative review for discussion purposes. The networked program used was designed to provide a challenge in math, science, reading, and writing. Comparing a group of students who received teacher-facilitated instruction on the program with a group that did not receive the facilitation (but did access the same technologies), the researchers found a significant effect of teacher facilitation on GPA and attendance (as measured at the end of the year). Students in the teacher-facilitated group showed significantly more gains in GPA and better attendance, although the authors do note that the gains in GPA were not enough — students in both groups entered the program failing academically and continued to fail following their exit. However, while the group with teacher-facilitated intervention showed an improvement of over six points in their GPA following the program, the group without teacher-facilitated intervention actually experienced a drop in GPA of two points over the year. The authors suggest that more efforts be made to ensure that disadvantaged youth and youth at risk of dropping out be given the opportunity to engage in *teacher-facilitated* remedial work, and that the exposure to the intervention continue for longer than one year.

Sankaran and Bui (2001) looked at 116 university students enrolled in a business course in one of two sections (face-to-face and Web-based lectures). The Web-based lectures, unlike some of the other papers, did not constitute a huge change from the traditional format, but rather were minimally adapted to be available for Web access. The authors measured performance on a pre-test and a post-test. They found no significant differences between the two groups on any of the measures, but did find, not surprisingly, that

students who were more motivated performed better in both formats. They use these findings to argue that instructors who are planning to offer their courses in Web format can do so with "minimum redesign," a finding contradicted by much of the other research. One aspect of this study worth mentioning is that they provide an analysis of performance based on ethnicity. They find no significant difference in performance across groups, offering a nice control for future research.

Schimmoeller (2003) offered an innovative approach to networked technology in the classroom. Investigating achievement and interest in science class for 103 grade 5 students, Schimmoeller combined solid research design with new technology. Science classrooms were fitted with Personal Response Systems, wherein students are expected to "vote" on certain questions as the class progresses. Students could take one of two classes and were in both on-line and traditional course formats, counterbalanced for groups across days. The authors were also interested in the role of gender, pointing out that there are traditional expectations associated with science, math, and computer science that may affect student behaviour. In addition, the experimenters noted that previous research has demonstrated that girls are less likely to raise their hands in class or ask questions, and they believed that the use of this technology might offset some of these gender biases. They measured the success of their intervention by comparing performance on a post-test. No differences were reported for either the students' gender or the presence of the intervention. Unfortunately, they did not have data on the number of questions answered in the two conditions as a function of gender, and so could not address one of the central foci of their investigation. This is a preliminary study and in some senses asks more questions than it answers, but still raises important issues and introduces a new and innovative approach to utilizing technology in the elementary school classroom.

Sumner and Hostetler (2002) studied 44 students in either an e-mail conference group or a traditional face-to-face meeting group and their communications on group projects. Students had to reach consensus on a variety of tasks and keep a log of the decision-making process. Although both groups scored similarly in terms of confidence in the group decision-making process, face-to-face groups reported higher levels of satisfaction with the process. The authors argue that this difference might be explained by relationship-building opportunities in the face-to-face group that are not present in the e-mail conferencing group, although we can compare opposite findings in similar situations — see discussion below. Interestingly, the e-mail conferencing group performed better (in terms of grades) on all the assignments, a result not reflected by student satisfaction ratings. The authors conclude that performance in e-mail conference decision-making is equivalent or better than in face-to-face groups, but that this is not perceived by the students in either of the two groups, suggesting a possible area of future intervention.

Thirunarayanan and Pérez-Prado (2002) compared the achievement of 60 university pre-service teachers who are learning to teach English as a Second Language (see Pérez-Prado and Thirunarayan above for the qualitative study). Through a series of measures, they evaluated learning in either a Web-based section or a traditional lecture-based section of the same course. They had a very clean experimental design, in that they

measured performance on a pre-test prior to the class, compared the means of the two sections, and then measured performance on a post-test and compared the two means again. They found significant differences between the groups on the pre-test (not described), with the lecture-instructed students demonstrating better performance. On the post-test (also not described), they found no significant differences between groups. They used these data to argue that the on-line students learned more during the class, although there is no significant difference in the change in scores between the two groups.

Wolverton and Wolverton (2003) studied 65 university students who had completed an on-line financial course and were surveyed regarding their learning and achievement in on-line courses versus face-to-face methods of instruction previously experienced. All students were in the on-line course and were also required to meet once a week for tutorial or seminar purposes. Students reported appreciating the opportunity to watch entire lectures or segments of lectures more than once, with a substantial proportion (37.3%) reporting that they reviewed difficult material with this strategy. A majority of students also reported finding the on-line format more interesting than traditional classroom instruction and would recommend it to a friend. They also reported that "exam scores were noticeably higher." These data are interesting and suggestive of the kind of benefits that can be derived from on-line learning — notably, the ability to set the pace and timing of learning to suit the individual. However, this study suffers from many of the same methodological weaknesses already reported, such as non-random assignment to the class and lack of an objective measure for a comparison group. As was done in other studies, the authors report on the large amount of time invested in the creation of the course (in excess of over 15 weeks full-time).

Yu and Yu (2001) studied 68 pre-service teachers in a Computers and Education class. The class was divided into two groups, one that used e-mail as a mode of communication for supplementary materials and one that did not. The authors focused on responses to a post-intervention survey of student attitude to e-mail use and the grades received by the two different groups, and claim that the use of e-mail improves student performance. The authors reported a statistically significant difference in academic performance between the two groups, with the e-mail group performing significantly better than the non–e-mail group. However, the authors did not have a pre-intervention measure of their performance. Although the participation in the classes was assigned randomly, not obtaining a baseline measure of performance to compare the two groups before the intervention lessens the causal strength of their argument. In terms of satisfaction with the course, no significant differences were found between the two groups. The authors offered an unconvincing argument to explain why this might be. They claimed that the number of e-mail viruses circulating dampened the enthusiasm of the e-mail intervention group (while presumably not affecting the non-e-mail group), thus resulting in no significant difference. The methodological errors and questionable arguments weaken the strength of the findings presented in this paper.

6.4 Discussion of findings raised by in-depth review

6.4.1 Diversity of the topics studied

The research included in the in-depth review was diverse in many ways including the variety of technologies, subjects, classrooms, and interventions investigated. The terms "on-line" and "networked" education clearly mean many different things to many different people. Technologies used in the studies ranged from hypertext homework tools to virtual (audio/visual) synchronous instruction. We were not surprised to find that the effectiveness of the technology seemed correlated with the extent of interactivity that the technology afforded the learners. The virtual synchronous classroom that allowed geographically diverse students to "attend" lectures together was much more positively reviewed than a course taught by e-mail lectures that were not supplemented by additional materials. Intranets also had different strengths than Internets; students using networked computers in a traditional class setting reported that they enjoyed the ability to interact as a group while actively solving problems in class. Of course, for distance education students who are not able to be present in a traditional classroom, the issues raised, both positive and negative, will be different. Even computer-mediated communication was broadly defined, with some courses involving e-mail, others e-mail and list-servs, and still others e-mail, listservs, and chat groups.

The type of interventions studied was equally diverse, ranging from entirely networked courses to single lectures, or traditional classes that required networked homework assignments and made use of on-line study tools. Some of the studies required the entire course to be networked and therefore looked at course achievement, while others investigated only single aspects of course performance, such as work on group papers or the ability to dissect argument style and structure in a debating course.

There was a very diverse assortment of subjects studied at many different levels, from science to writing in elementary school to English literature and physics in secondary schools. At the university level the diversity was even greater, including business, psychology, health education, physics, composition, Spanish, social work, teaching ESL, and debating and public speaking. The different subject matters allow for greater generalizability of themes and are a strength of the review. By generalizability, we mean that if something is true in a variety of different environments, it is more likely to be generally true and not just a function of a specific subject. For example, students in many of the university studies reported enjoying the flexibility of networked education and the ability to control the pace of their learning. As a common theme emerging across so many diverse disciplines, it seems very likely that this is an aspect of networked education that is generally true. On the other hand, the diversity of the subject matter also allows us to investigate particular weaknesses of networked education. For example, three different papers on social work and teacher education highlight the concern that subject matters requiring an affective component or empathetic reaction might not be as suitable to on-line methods of instruction.

6.4.2 General themes

The 25 studies in the in-depth narrative review present a mixed picture of the effectiveness of on-line learning. Two studies (Sumner and Hostetler, 2002; Johnson, Aragon, Shaik, and Palma-Rivas, 2000) report significantly higher levels of student satisfaction with the traditional format as compared to on-line formats. One study showed a positive effect of traditional instruction in terms of grades. Looking at group work and collaborative learning and comparing traditional groups to e-mail groups, Barile and Durso (2001) found that the traditional group performed significantly better, leading the authors to conclude that e-mail is not an effective method of group collaboration.

Five studies showed a reversed difference, with on-line instruction having a significantly more positive effect than traditional classrooms. However, three of those studies were methodologically unsound and were given a 'low' weight in the review (Buchanan, 2000; Grabe and Sigler, 2002; Yu and Yu, 2001; see section 6.3.4 for a fuller discussion). Of the remaining two studies, one (Hertz-Lazarowitz and Bar-Natan, 2002) argued that computer-mediated writing is just as effective as using computers plus traditional instruction techniques for elementary school children, a finding the authors argued contradicts widely held beliefs about the ineffectiveness of computers in teaching language arts. The other study (Bain, Huss, and Kwong, 2000) illustrated that the use of a hypertext tool for homework significantly improved scores compared to traditional homework, although the findings are confounded in that the tool took more time and thus could be a reflection of time spent increasing scores on homework more so than the tool itself.

In general, the number of studies purporting to show an effect of one method over the other (traditional vs. on-line) is small and many of the claims being made on both sides are mitigated by methodological concerns or other factors. Certainly there is nothing present in the studies reviewed here to allow us to make strong statements one way or the other.

Of the 13 studies that showed no significant difference between forms of instruction, at least three are methodologically unsound. However, there are a number of studies here, from different subject backgrounds and different technologies, that are strong and offer compelling tests of effectiveness. Some of the themes that emerge from these and the qualitative research include advantages and disadvantages of both forms of instruction.

Several advantages were attributed to traditional learning: interacting with the instructor and other students, the structure that a particular class time gave to the day, and the ability to get feedback and have questions answered in the "here and now." Disadvantages include the obvious lack of flexibility in course time and location.

Likewise, positives were attributed to networked learning: the flexibility to learn where and when one wishes at one's own pace, and the ability to review material presented. Three studies expressed concern about the use of on-line education for teaching courses with an affective component, such as social work or teacher education. Concerns about collaboration in on-line environments, raised above by Barkhi and Brozovsky (2000) are

mitigated by the findings of Barile and Durso (2002) and two studies investigating public speaking (Clark and Jones, 2001) and debating (Marttunen and Laurinen, 2001) which all showed no significant difference between on-line and traditional instruction groups in student achievement.

Two additional themes should also be highlighted. First, there is a concern about the high levels of attrition by networked students compared to students in traditional environments (Clark and Jones, 2001; Dutton, Dutton, and Perry, 2002). Indeed, Dutton et al. report a 20% greater decrease in completion rates for their on-line section in contrast to the face-to-face section. This is a sobering statistic that merits further investigation. Related to this point is the second concern that, although measures of achievement were often not significantly different between groups, face-to-face groups reported higher satisfaction and perceived achievement ratings (Sumner and Hostetler, 2002; Johnson, Aragon, Shaik, and Palma-Rivas, 2000). If students enjoy themselves more and perceive themselves as performing better in a traditional environment, it seems likely that retention will be enhanced. If the opposite is true for on-line sections, regardless of objective performance measures, students might be more inclined to withdraw from the course.

6.4.3 Gaps

Despite the range of subjects and technologies studied, there are also gaps that can be identified in the reviewed research. One of the most salient points comes from a quick look at the study characteristics outlined in Section 6.1. By far the greatest proportion of research has been conducted at the university level. We were only able to identify two studies on secondary students (one of which did not directly compare the effectiveness of on-line versus traditional instruction and was only included because it focused on disruptive students) and two devoted to elementary students. Clearly, more work needs to be done with elementary and secondary students as there are very specific issues related to their development and education for which we currently have no answers.

Another gap in the research involves issues of inclusion, such as the role of gender, ethnicity, urban/rural, and disabled/disadvantaged populations. Very few of the studies looked at gender, although several raised gender as a topic of interest and highlighted important issues regarding the possible role of technology in overcoming gender bias in class participation and interaction. None of the research looked at special populations such as students with learning disabilities or behavioural problems — with the exception of Powell, Aeby and Carpenter-Aeby (2003), which, as explained above, did not actually have a comparison group that addressed our research questions. We also found no research that investigated the role of socio-economic status, although many of the papers discussed concerns related to accessibility and computer ownership. And, sadly, we found absolutely no research that investigated the role of urban vs. rural location and networked education, although, of course, that is one of the primary justifications for distributed electronic learning or distance education.

All the gaps raised here deserve further scrutiny and will be addressed in the discussion and recommendation sections of this report.

6.4.4 Methodological issues

Only 10 of the 25 studies included in the in-depth review were not seriously flawed, a sobering statistic given the constraints that went into selecting them for the review. Studies were commonly flawed either in design, statistics, or interpretation. An example of poor design is given in Buchanan (2000). In this study, 214 university students were given the option of using a set of on-line study tools. Because use of the tools was voluntary, the researchers ended up with dramatically different numbers of students in each condition: 16 students used the study tools and 198 did not. Not surprisingly, they found that the students who used the study aids scored significantly higher on some measures of learning. Unfortunately, as use of the tools was voluntary, it is not hard to imagine a scenario where the best and most motivated students used all available means to study for the exam, and thus performed better – not because of the study tools themselves, but because these were the best and brightest students. A similar flaw was evidenced in the work of Grabe and Sigler (2002).

It is noteworthy to highlight the one qualitative study that was included in the in-depth review. Pérez-Prado and Thirunarayanan (2002) offer a comparison of student perspectives of on-line versus face-to-face teaching. By asking students to keep a journal of their experiences and coding both the journal and interviews, the authors were able to obtain perspectives on the difficulties and benefits of Web-based instruction using qualitative methodology. It is important to note that, because they had a comparison group of students who also kept journals in a traditional class format, the authors were able to make stronger claims about the thoughts and opinions expressed by the on-line group. This qualitative study provides clear and strong arguments commensurate in quality to any of the quantitative studies included in the review. Unfortunately, none of the other qualitative research included a comparison group and thus did not meet our criterion for inclusion in the in-depth review.

7. Meta-Analysis: Methodology and Results

7.1. Study characteristics

As previously outlined in the narrative review, all studies that met our inclusion criteria and had a comparison group were selected for in-depth review. Of the 37 studies that met these criteria, 12 reported enough statistical information to allow for inclusion into a meta-analysis. Of the 12 studies, 10 were from the USA, one was from Canada, and one was from Greece. In terms of the educational setting, the same pattern observed in the narrative review was repeated here: 10 of the 12 studies (85%) were conducted at the university level. Two studies were conducted in secondary schools, and no studies investigated elementary school students. Even in this selective sample, methodological quality was questionable. Only 2 of the 12 studies had random assignment of participants to groups and none of them had blind raters.

7.2 Table of studies

Author, Date, and Country	**Study Type**	**Aim**	**What was studied?**	**How?** Note: ES = Effect Size
Bain, Huss, Kwong (2000) USA	quantitative (survey)	Evaluate a hypertext tool for teaching English literature to secondary school students	• sample: 39 grade 11 English literature students in three classes • achievement • attitude	Attitude survey, weekly tests (repeated measures) ANOVAs on tests; descriptive analysis of survey data ES: calculable
Barkhi, Brozovsky (2000) USA	quantitative (survey)	Analyze the dynamics of a virtual classroom compared to a traditional classroom	• sample: 62 university students (course unknown) • perception of learning experience • achievement on tests	Attitude survey, outcomes on midterm and finals ANOVAs on tests, survey data ES: YES
Collins (2000) Canada	quantitative (survey)	Compare Web and lecture versions of a second year non-major biology course	• sample: 173 university students enrolled in one of two sections (Web or face-to-face) of a biology course • comparing performance, satisfaction of students as a function of instructional format	attitudinal survey; final marks of several sections t-tests, ANOVA ES: calculable
Dufresne, Mestre, Hart, Rath (2002) USA	quantitative (quasi-experimental)	Determine effect of Web-based homework on test performance in introductory physics courses	• sample: unknown *but very large* number of university students in 15 physics classes • achievement as a function of either Web-based homework assignments or traditional assignments	Exam scores, SAT scores, homework scores t-tests ES: calculable
Faux, Black-Hughes (2000) USA	quantitative	Investigate performance of Social Work students in a traditional classroom vs. an Internet classroom	• sample: 33 university Social Work students in one of three classrooms (traditional, Internet, and traditional + Internet). • Performance and knowledge of the History of Social Work.	pre- and post-tests Descriptive statistics, ANOVAS ES: calculable

Author, Date, and Country	Study Type	Aim	What was studied?	How? Note: ES = Effect Size
Hartzoulakis (2002) Greece	quantitative	Investigate computer-mediated discussion in the teaching of ESL	• sample: 24 Greek secondary school students in an English course • choosing participants whose English was poorest; investigating learning in either a Web-based or traditional format	Pre- and post-tests of performance Descriptive statistics; ANCOVAs, ANOVAs ES: calculable
Maki, Maki, Patterson, Whittaker (2000) USA	quantitative (survey)	Evaluate learning and satisfaction of on-line versus traditional introductory Psychology courses	• sample: 218 Introductory Psychology students in either an on-line or face-to-face format • performance and satisfaction as a function of course format (across years)	attitudinal surveys; comparisons of final marks and GRE questions ANOVA ES: calculable
Sankaran, Sankaran, Bui (2000) USA	quantitative (survey)	Comparing the effect of student attitude to course format on learning performance	• sample: 116 university business computer course students (in one of two groups) • performance on tests and attitudes to computers	attitudinal survey; final test scores and midterm scores t-tests; descriptive statistics ES: calculable
Schoenfeld-Tacher, McConnell, Graham (2001) USA	quantitative and qualitative	Compare the effects of on-line vs. traditional delivery of a science course	• sample: 44 university students in a Histology course (on-line and face-to-face sections) • performance on exams and quality of interactions among students	pre- and post-tests; chat transcripts, and lectures ANCOVA, ANOVA ES: calculable
Tuckman (2002) USA	quantitative	Evaluate a hybrid instructional model combining Web-based and classroom components	• sample: 452 university 'study skills' students in three groups (Web-based, face-to-face, control) • achievement in course based on type of instruction	pre- and post-intervention GPA ANCOVAs ES: calculable

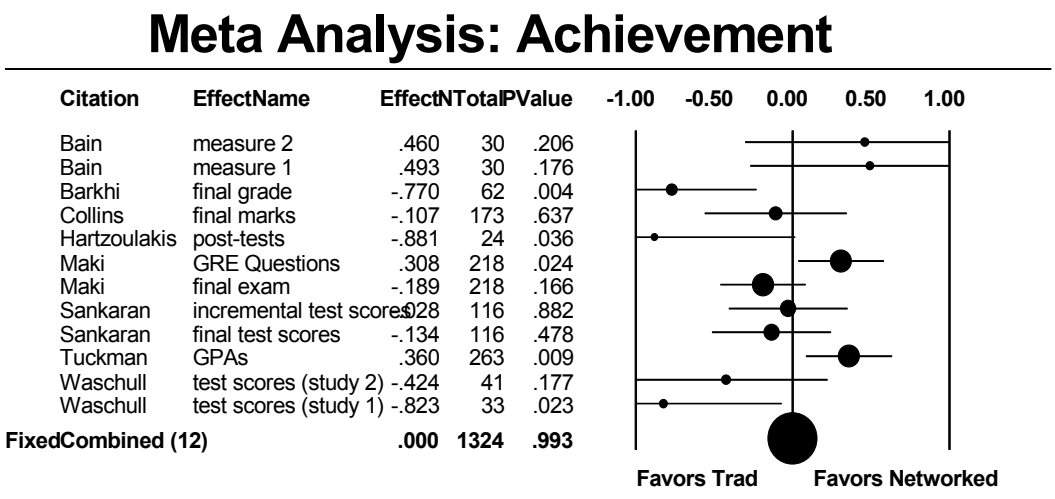
Author, Date, and Country	Study Type	Aim	What was studied?	How? Note: ES = Effect Size
Wang, Newlin (2000) USA	quantitative (survey)	Characterize students who succeed at Web-based Psychology classes	• sample: 105 university Statistics for Psychology students in two groups (Web-based and face-to-face lectures) • Achievement, attitudes, and work habits of students in different types of course	Attitudinal survey, study habits survey, on-line course activity, and course grades descriptive statistics, ANOVAs, t-tests, correlations ES: calculable?
Waschull (2001) USA	quantitative (survey)	Compare on-line versus traditional formats of Psychology courses for attrition and performance	• sample: 74 students (in two studies) enrolled in an introductory Psychology course • Retention rates across groups, performance, and satisfaction	course evaluation, test scores, class roll chi-square, Bonferroni adjustment ES: calculable

7.3 Meta-Analysis: Results

Four studies were excluded from the meta-analysis for fatal flaws in either methodology or statistical analyses. Dufresne, Mestre, Hart, and Rath (2002) do not present an appropriate statistic to enter into our meta-analysis, and those statistics that were presented were questionable. They do not provide aggregate means or standard deviations across classes, and although they do provide gain scores, they do not include a standard deviation. Similarly Faux and Black-Hughes (2000) make numerous statistical errors in their reporting, including labelling a p value an F and reporting the same thing as a standard deviation in one sentence and a standard error in a table. Because of these errors, we do not have confidence in their statistics and did not feel it was appropriate to include them. Schoenfeld-Tacher, McConnell, and Graham (2001) offer appropriate statistical analyses but have a design flaw that confounds their results. Their experimental groups are given either face-to-face lectures or Internet lectures, but both groups are required to take assessments on the Internet. It is not surprising that the face-to-face group, which purposely had no Internet experience prior to testing, would perform more poorly on the Internet tests. Also, it is unclear if the results are confounded by total time spent; on-line students appear to receive almost double class/discussion time as lecture students, but unfortunately there was no way to verify this as time per section was not presented in the methodology. And finally, Wang and Newlin (2000) also present questionable statistics. They report a mean of 89.8% (SD: 0.9) for their face-to-face section, and a mean of 85.7% (SD: 1.1) for their on-line section. This is not a normal distribution and as such is not an appropriate sample on which to perform statistical analysis. We suspect that they may have confused SE with SD, but there is no way to discern if this was the case or not.

The other 8 studies present an interesting story that complements the findings of our narrative review. Overall, no clear difference is observed (see Figure 1) when comparing the effect of networked versus traditional education.

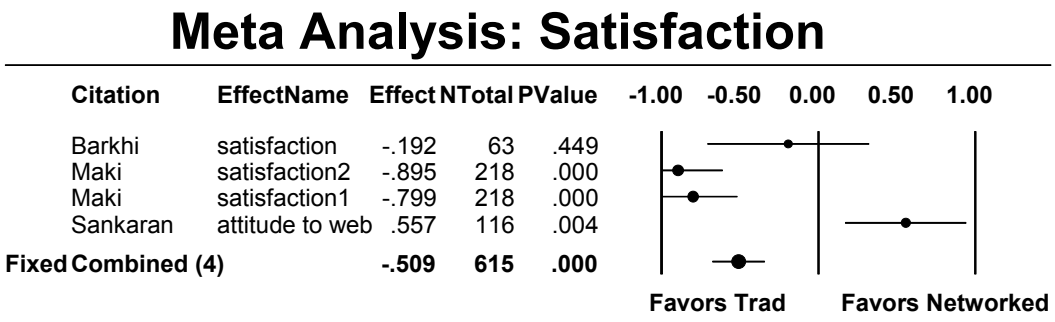
Figure 1. Meta-Analysis: Achievement



As Figure 1 demonstrates, there is no overall difference between the different modes of instruction. Although some studies find the traditional group performing better than the networked group, others have the opposite result. Clearly, based on the literature reviewed here, there is no significant difference between the two methods of instruction in terms of academic achievement (grades, test scores, etc.).

This finding pertains to an issue often raised in terms of technology and education: that the standard methods of assessment used for traditional teaching are not appropriate for the kind of learning being done in a technology-rich environment. If this is true, finding no significant differences in terms of academic achievement (assessed using a traditional conception of knowledge mastery and learning) merely reflects the inability to assess the changes that are being produced by technology. It seems, then, that it might be more appropriate to look at other measurements of "success" and "failure." One option is to look at measures of student satisfaction and sense of mastery with their learning, independent of test scores. Although not all of the research offered relevant measures with appropriate statistical information, three of the studies did provide this data. Accordingly, we conducted a second analysis that focused on student ratings of satisfaction (see Figure 2).

Figure 2. Meta-Analysis: Satisfaction



Although our sample size was small and, thus, our results should be treated with caution, the meta-analysis is still suggestive. A significant difference was found in satisfaction levels favouring the traditional methods of instruction ($N = 615$; $p < .0001$). The weighted mean effect size was $g = -.509$. The 95% confidence interval for the standard error was -.675 to -.344, indicating a significantly positive effect for traditional classrooms as compared to on-line classrooms. Interestingly, the very studies that show a strong effect for traditional instruction are also the same studies that showed higher achievement in the on-line sections in the previous analysis (Barkhi et al., 2000; Maki et al., 2000). This echoes findings of our narrative review, in which measures of achievement were often not related (or inversely related) to measures of student satisfaction. Also of note is that all of the measures showed the effect in the same direction except for Sankaran et al. (2000). Sankaran's outcome measure is not actually a measure of satisfaction with the course, but rather a measure of students' attitudes toward the Internet. Therefore, it is not surprising that they found that students who choose on-line sections have a more positive attitude toward the Internet, and caution that standard

assessments of students in non-randomly assigned sections are likely to be influenced by self-selection. As an addendum, it was disappointing to see that although some authors commented on why measures of achievement might not be appropriate, very few provided any other suggestions. An exception is Collins (2000), who offers a strong analysis of achievement and also observes that "simply comparing student performance in Web and traditional courses is not the best way of deciding on the success of such new approaches. However, such a comparison should be considered as a first step in the evaluation" (p.26). At the same time, however, he provides an evaluation of student satisfaction for Web-based courses only, omitting (or not collecting) the same information from lecture-based courses, thus thwarting meaningful comparison.

8. Meta Analysis: Discussion

8.1 General themes

Our meta-analysis looks at the effectiveness of networked versus traditional forms of instruction by focussing on two measures: achievement measures, generally conceptualized as grades, and satisfaction ratings. We find no difference between forms of instruction on grades and achievement indices, and a significantly negative effect of on-line instruction on student satisfaction. These data are comparable to the themes that were raised in our narrative review, and are interesting results in light of the meta-analysis performed by Bernard, Lou, and Abrami (2003), which finds a small but significant positive effect of interactive distance education over traditional education for academic outcomes, and a negative effect for retention rate and student attitude toward subject matter. Similarly, in their meta-analysis, Allen, Bourhis, Burrell, and Mabry (2002) find a student preference for traditional instruction, but no difference in satisfaction levels between the methods of instruction. All of these sources (our meta-analysis, our narrative review, and two other meta-analyses) point to student satisfaction or preference for traditional instruction which does not seem to be correlated to actual performance. Students appear to perceive networked instruction less favourably than traditional instruction, regardless of their test scores.

This finding is very probably related to an issue that is raised in both our meta-analysis (Maki et al., 2000) and narrative review (Clark and Jones, 2001; Dutton, Dutton, and Perry, 2002), and is echoed in the work of Bernard, Lou, and Abrami (2003) — that networked courses have lower retention rates than traditional courses. It seems reasonable to argue that, if students are not enjoying their experience and feel like they are achieving less in a networked environment, they will be more likely to discontinue the course, regardless of their actual academic achievement. This is an important aspect for policy and planning, and offers a suggestion for future work.

Another theme that echoes the findings of our narrative review is the argument that networked learning is best suited to particular forms of instruction. Three papers in our narrative review raise concerns about the use of networked education in the teaching of courses with an affective or empathetic component, such as social work or teacher education. Similarly, in the meta-analysis Barkhi and Brozovsky (2000) argue that virtual classrooms support particular types of learning best, that is, ones that utilize independent learning. For courses requiring group work and conflict resolution, the authors propose investigating other options in a virtual setting.

8.2 Gaps

Again, the gaps of our meta-analysis echo those observed in our narrative review. A quick look at the study characteristics makes it clear that most of the research is conducted in the United States. The Canadian context is unique and deserves research that targets issues pertinent to Canada's population, such as linguistic diversity and the urban/rural divide. Similarly, most of the research is conducted at the university level. In our meta-analysis, no studies were conducted with elementary-level students and only two were done with secondary level students. This is a serious gap in that many of the issues faced by children and adolescents in learning environments are incommensurate with those faced by adults. We cannot generalize from adult learning strategies and performance to school-aged children without running the risk of being fundamentally misguided in our assumptions. The lack of research on K–12 students is troubling, particularly given the findings of Cavanaugh's (2001) meta-analysis of the effectiveness of interactive distance education technologies in K–12 learning. Although she finds a small positive effect in favour of distance education on a number of dimensions, she is not looking specifically at networked education and does not specifically address achievement outcomes. However, there are important findings specific to this population: she finds a positive effect for small group size and programs of short duration, and a negative effect for primary instruction through distance.

8.3 Methodological issues

Out of the 232 studies with which we began, we were only able to find 12 studies that met our criteria and contained enough statistical information for inclusion in the meta-analysis. None of the researchers we contacted to request more statistical information responded to our request. In addition, of the 12 studies that did meet our criteria for inclusion in the meta-analysis, 4 suffered from such serious statistical or design flaws that they were excluded from the analysis. This does not speak highly of the research done in this area, and once again reiterates a point raised in our narrative review.

That being said, there are several examples of innovative and sound study methodology in this sample. Bain, Huss, and Kwong (2000) studied grade 11 English Literature students' achievement and attitudes toward a hypertext tool in a repeated measures design. The study lasted four weeks: in weeks 1 and 3, students were given traditional lectures and independent reading for homework (baseline). In weeks 2 and 4, the classroom teaching component remained the same but students were required to use the hypertext tool as a replacement for traditional homework. At the end of each week, students were given a test that assessed comprehension of the preceding week's instruction. In addition, an attitude survey about the students' experiences with the study tools was administered at the end of the fourth week. The results indicated that the hypertext tool was a more effective method of study than the traditional method. However, students' attitudes toward the tool were ambivalent; although some reported enjoying the process, many others did not. Overall, students spent more time on their homework in the intervention weeks than during the baseline weeks, a fact the

researchers point out might explain both their higher achievement and lower appreciation for the tool. It is interesting to note that researchers used a tool with which the students were all familiar, one that had been required in a previous unit. They argue that this had the effect of decreasing any potential novelty effect, thus mirroring, to some extent, a more real-world situation.

This study is a useful illustration because it offers the possibility of conducting a controlled experiment in a real-world educational setting. By taking care to avoid novelty effects, the authors also add validity to their findings and can make stronger arguments about long-term reactions to technology in the classroom. Of course, this is not an option for students who are physically removed from a given geographical location, and we acknowledge that not all research on networked education can, or should, be conducted in a similar manner.

9. Conclusions

The research on the effectiveness of networked and on-line technologies is still in the developmental stage. Our review raises many issues about the quality and rigour of research conducted in this area. There is a clear need for more systematic research on the effect of networked and on-line technologies, particularly with a focus on children and adolescents, women/girls, and people from marginalized populations.

Although much of the experimental research is limited in its quality, there has been an impressive series of correlational studies conducted on very large populations. SITES in Canada, ImpaCT2 in the UK, and Hakkarainen and Ilomaki (2000) in Finland have all conducted country-wide investigations into the effect of information and communications technologies (ICTs) on student achievement and attitudes. Statistics Canada has also recently completed an analysis of country-wide data gathered from school-age students, looking at computer use and academic achievement as a function of a number of socio-economic variables (Looker and Thiessen, 2003). This work is particularly relevant to our discussion of gaps in the research. Their data illustrate that traditional divides still exist, with females, students from families of lower socio-economic status and lower parental education showing less confidence in their computer use and abilities. Of special note is their analysis of the urban/rural divide and use of ICTs. Overall, rural students and rural schools seem disadvantaged in various ways compared to their urban counterparts. Rural youth are less likely to have access to computers in their homes, and even more limited access to the Internet; the computer support and educational software in their schools is less specialized; and their teachers have less access to ICT support. Although Looker et al. argue that rural schools seem to be making up for the lack of computers in the homes, this argument should be approached with caution. Ravitz, Mergendoller, and Rush (2003) recently demonstrated that higher in-school computer use is *inversely* related to academic achievement. Instead, home use of computers is most strongly correlated with higher academic achievement, a finding echoed by the ImpaCT2 study (Harrison et al., 2003).

These studies, while intriguing, are correlational by nature and so cannot offer the explanatory power of empirical research. As spelled out in the discussions of both the narrative review and the meta-analysis, there is a dire need for rigorous research in this area in order to provide a basis for evidence-based policy formation.

10. Recommendations

Federal, provincial and municipal governments, and public agencies such as school boards, colleges and universities, non-governmental agencies and private sector organizations have made significant investments in information and communication technologies. The investments include the development of infrastructure to support access to the Internet, organizational intra-nets, general and special purpose portals, as well as software. While there are few reliable estimates of the magnitude of the investment, there is little doubt that the investments have been costly.

In an effort to address the absence of systematic knowledge about "the expansion of knowledge in all fields, the proliferation of communications technologies, and the globalization of markets for goods, services and ideas . . . ", the Social Science and Humanities Research Council of Canada fosters research under the ambit of *Initiative on the New Economy* (see http://www.sshrc.ca/web/apply/background/ine_about_e.asp. Though such effort is praiseworthy, much more systematic and programmatic research is needed before strong conclusions can be drawn about the effectiveness and efficiency of on-line and networked learning.

> *Federal, provincial and municipal governments, public agencies, non-governmental agencies and private sector organizations should reserve a proportion of their significant investments in on-line and networked learning for well-conceived, well-executed programmatic studies of their efficiency and effectiveness.*

In addition to the relatively spare investment in research, much of the research that has been conducted has been undertaken in the United States. While the similarities between the United States and Canada are likely greater than the differences, the differences are important. Canada's dependence on immigration and the need to integrate persons from diverse linguistic backgrounds into the fabric of the nation are but two of the demographic differences that distinguish Canada from her southern neighbour. Canadian industry exhibits lower productive efficiency than some other nations, and there are large variations among provinces (see, for example, http://www.statcan.ca/english/sdds/1402.htm, http://www.statcan.ca/Daily/English/010824/d010824a.htm). The impact of such differences on the use of information and communication technologies in general and on-line and networked learning in particular deserve much closer scrutiny.

Research devoted to the effectiveness and efficiency of on-line and networked learning should take into account contextual variables that distinguish Canada from other nations and are likely to influence the effectiveness and efficiency of on-line and networked learning.

While we are very much committed to the value of basic research free from constraint as to the nature and direction that such research should take, the research that is needed is applied and should be programmatic in nature. To that end, it would be advantageous to establish research advisory committees to assist in the identification of priorities and the adjudication of research proposals relevant to the domain (university, elementary and secondary education, industrial or service sector) with which the sponsoring agencies are concerned.

Care should be taken in empanelling a research advisory committee to avoid the polarization that has occurred. We have observed that some members of the educational research community have committed themselves to the use of particular research methods in advance of formulating a particular research question — making the selection of research methodologies akin to choosing up sides in a hockey game. We believe that the selection of research methodologies must be subordinate to and occur after the formulation of a particular research question, not prior. Different questions require different methodologies. Rich, qualitative methodologies are particularly helpful in acquiring a detailed understanding of processes. Randomized experiments are useful for hypothesis-testing.

Sectoral research advisory committees should be established to assist in the identification of priorities and the adjudication of research proposals relevant to the sectors with which the sponsoring agency is concerned.

At present, there is an imbalance in the amount of attention devoted to on-line and networked learning. Approximately 77% of the studies we identified were conducted in university settings, with the remaining effort distributed among elementary and secondary schools, workplaces, and community centres. There is a need for more and better research in all sectors, but especially in workplaces, secondary, and elementary schools. There is also a need for research in other types of education where adults make up the student population and for special populations such as aboriginals. Given that the nature of the activities conducted in workplaces is so varied, it will likely take significant effort and resources to adequately investigate the efficacy and efficiency of on-line and networked learning in the workplace. There is a significant body of research to support the claims that (1) children and adults learn differently, and (2) adults with varying prior experiences learn differently. Thus, it would be a mistake to make inferences from studies of adult university learners to other adult learners or to elementary and secondary students. For these reasons, and because of the significant investment in information and communication technologies in the elementary and secondary schools, we believe that priority should be given to research with children and youth of school age.

Significant attention should be devoted to research in all domains, but especially to the use of networked and on-line learning in elementary and secondary schools, the workplace, other types of education where adults make up the student population, and special populations.

Internet technologies have been favoured in the research literature. Our estimate is that such studies account for approximately 60% of the recent available research. The remaining 40% is distributed among studies of networked environments, tele-teaching, and other computer-mediated technologies. These latter technologies deserve more attention until such time as it is possible to make robust claims about their effectiveness and efficiency.

Significant attention should be paid to research involving all available technologies, but especially to networked environments, tele-teaching, and other computer-mediated technologies.

Less than a third of the studies devoted to on-line and networked learning that we identified and reviewed made use of control or comparison groups. We regard this as a significant shortcoming in the research. Indeed, we favour the use of randomized experiments where it is feasible to do so.

There are a number of objections to randomized experiments, including that they (a) suppose an oversimplified theory of causation, (b) suppose an oversimplified epistemology, (c) are not suited to complex organizations, (d) are premature until a good theory of the program and its mediating processes has been developed, (e) are politically infeasible, (f) have been tried and failed, (g) entail trade-offs not worth making, (h) are unethical, and (i) are not needed because alternatives are available that are as good.[3] We are nonetheless persuaded by the argument advanced by Cook and Payne that random assignment provides (a) a logically more valid causal counterfactual than any of its plausible alternatives, (b) a more efficient counterfactual, and (c) a counterfactual that is more credible.[4] But we are also cognizant that it is necessary to "extend the measurement, sampling, and analysis frameworks of experiments" in order to explore an interesting range of questions.

[3] Cook, T.D. & M.R. Payne. (2002) Objecting to the objections to using random assignment in educational research. In *Evidence matters: Randomized trials in education research*, pp150–178. (Eds. Mosteller, F. and R. Boruch, Brockings Institution Press)

[4] *Ibid*. p. 174

Investigations of on-line and networked learning, tele-teaching, e-mail, and other computer-mediated technologies should make more extensive use of randomized experiments [randomized control trials] where the experimenter is blind to the conditions to which participants have been assigned.

Among the many justifications offered for the significant investment in the development of information and communications technologies — including networked and on-line learning — are that they (a) provide access to persons for whom access is difficult or impossible (e.g., rural and remote learners; persons with disabilities; persons whose schedules or commitments do not permit them to attend schools or postsecondary institutions); (b) reduce the likelihood of inequalities developing between those with access and those without access; (c) promote achievement for students whose learning styles and needs differ from those who benefit from conventional forms of instruction; (d) are necessary for ensuring Canada's position in the "knowledge economy"; (e) promote gender equality for women. Given the prominence they receive, we expected that these issues would be investigated. The relative absence of research addressing such justifications surprised us.

Investigations of on-line and networked learning, tele-teaching, e-mail, and other computer-mediated technologies should attempt to determine whether the claims made for such technologies — that they (a) provide access to persons for whom access is difficult or impossible (e.g., rural and remote learners; persons with disabilities; persons whose schedules or commitments do not permit them to attend schools or postsecondary institutions); (b) reduce the likelihood of inequalities developing between those with access and those without access; (c) promote achievement for students whose learning styles and needs differ from those who benefit from conventional forms of instruction; (d) are necessary for ensuring Canada's position in the "knowledge economy"; (e) promote gender equality for women — are justified.

We believe that conventional approaches to funding, such as those used by the Social Sciences and Humanities Research Council of Canada, are not appropriate for the programmatic research that is needed in the field. Establishing the domain as a funding priority and eliciting proposals for research will not address the need for programmatic research. The gaps in knowledge are too numerous and, judging by the research we have reviewed, the quality of the research too poor to advance understanding of an area that has received significant public and private investment. The situation is unlikely to improve without significant financial commitment to research and guidance to researchers about priority research questions within the domain and about the kinds of designs that are needed to address the questions identified. (See also: CMEC On-Line Working Group (July 2001) http://www.cmec.ca/postsec/on-lineLearningEN.pdf; the Advisory Committee on On-line Learning (February 2001) http://www.cmec.ca/postsec/evolution.en.pdf)

A significant financial commitment should be made to funding research about the effects, effectiveness, and efficiency of on-line and networked learning. Only the Government of Canada has sufficient resources to provide the financial commitment necessary to advance knowledge in this domain.

A research advisory committee should be established to recommend priorities for funding and to provide guidance to researchers about priority research questions within the domain.

Research capacity should be developed among persons interested in networked and on-line learning. Such capacity should include attention to research designs and methods needed to address the questions identified by the research advisory committee.

References

Adrianson, L. (2001). Gender and computer-mediated communication: Group processes in problem solving. *Computers in Human Behavior, 17*, 71–94.

Aldridge, J. M. *Grade 11 online nuclear physics course: Investigating the learning environment and student outcomes.* Paper presented at the American Educational Research Association, (2003). Chicago.

Allen, M., Bourhis, J., Burrell, N., & Mabry, E. (2002). Comparing student satisfaction with distance education to traditional classrooms in higher education: A meta-analysis. *The American Journal of Distance Education, 16*(2), 83–97.

Andrews, R., Burn, A., Leach, J., Locke, T., Low, G., & Torgerson, C. (2002). A systematic review of the impact of networked ICT on 5–16 year olds' literacy in English (EPPI-Centre Review). In: *Research Evidence in Education Library*. Issue 1. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

Bain, A., Huss, P., & Kwong, H. (2000). The evaluation of a hypertext discussion tool for teaching English literature to secondary school students. *Educational Computing Research, 23*(2), 203-216.

Barile, A. L., & Durso, F. T. (2002). Computer-mediated communication in collaborative writing. *Computers in Human Behavior, 18*(2), 173–190.

Barkhi, R., & Brozovsky, J. (2000). An analysis of the dynamics of a distance course. *Journal of Education Technology Systems, 28*(4), 311–325.

Bernard, R. M., Lou, Y., & Abrami, P. (with research team Wozney, L., Borokhovski, E., Wallet, P. A., Wade, A., & Fiset, M.). *How does distance education compare to classroom instruction? A meta-analysis of the empirical literature.* Paper presented at AERA 2003, Chicago.

Buchanan, T. (2000). The efficacy of a world wide web. *Journal of Computer Assisted Learning, 16*, 193–200.*Computing Research, 23*(2), 203–216.

Cavanaugh, C. S. (2001). The effectiveness of interactive distance education technologies in K–12 learning: A meta-analysis. *International Journal of Educational Telecommunications, 7*(1), 73–88.

Clark, R. A., & Jones, D. (2001). A comparison of traditional and online formats in a public speaking course. *Communication Education, 50*(2), 109–124.

Collins, M. (2000). Comparing Web, correspondence, and lecture versions of a second year non-major biology course. *British Journal of Educational Technology, 31*(1), 21–27.

Cooper, L. (2001). A comparison of online and traditional computer applications classes. *THE Journal, 28*(8), 52-56.

Dewhurst, D. G., Macleod, H. A., & Norris, T. A. M. (2000). Independent student learning aided by computers: An acceptable alternative to lectures? *Computers & Education, 35*, 223–241.

Dominguez, P., & Riley, D. (2001). Assessing distance education courses and discipline differences in their effectiveness. *Journal of Instructional Psychology, 28*(1), 15–19.

Dufresne, R., Mestre, J., Hart, D. M., & Rath, K. A. (2002). The effect of Web-based homework on test performance in large enrolment introductory physics courses. *Journal of Computers in Mathematics and Science Teaching, 21*(3), 229–251.

Dutton, J., Dutton, M., & Perry, J. (2002). How do online students differ from lecture students? *Journal of Asynchronous Learning Networks, 6*(1), 1–20

Faux, T. L., & Black-Hughes, C. (2000). A comparison of using the Internet versus lectures to teach social work history. *Research on Social Work Practice, 10*(4), 454–466.

Gough, D.A., Kiwan, D., Sutcliffe, S., Simpson, D., & Houghton, N. (2003). A systematic map and synthesis review of the effectiveness of personal development planning for improving student learning. Retrieved May 12, 2003, from http://www.ltsn.ac.uk/embedded_object.asp?id=18426&prompt=yes&filename=PDP038. The full report will be published on the Generic Centre Web site following peer review.

Grabe, M., & Sigler, E. (2002). Studying online: Evaluation of an online study environment. *Computers & Education, 38*(4), 375–383.

Green, R., & Gennteman, K. (2001). *Technology in the curriculum: An assessment of the impact of online courses*: George Mason University. Retrieved May 12, 2003, from http://assessment.gmu.edu/reports/Eng302Report.pdf.

Hakkarainen, K., & Ilomaki, L. (2000). Students' skills and practices in using ICT: results of a national assessment in Finland. *Computers & Education, 34*, 103–117.

Harrison, C., Comber, C., Fisher, T., Haw, K., Lewin, C., Lunzer, E., McFarlane, A., Mavers, D., Scrimshaw, P., Somekh, B., & Watling, R. (2003). *ImpaCT2: The impact of information and communication technologies on student learning and achievement* (ICT in Schools Research and Evaluation Series No. 7): DfES.

Hartzoulakis, V. (2002). *Investigation of computer mediated discussion in a Greek state lykeio.* Unpublished Master's Thesis, Athens.

Hertz-Lazarowitz, R., & Bar-Natan, I. (2002). Writing development of Arab and Jewish students using cooperative learning (CL) and computer-mediated communication (CMC). *Computers & Education, 39*(1), 19–36.

Johnson, M. M., & Huff, M. T. (2000). Students' use of computer-mediated communication in a distance education course. *Research on Social Work Practice, 10*(4), 519–532.

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

Lockyer, L., Patterson, J., & Harper, B. (2001). ICT in higher education: Evaluating outcomes for health education. *Journal of Computer-Assisted Learning, 17*(3), 275–283.

Looker, D.E., & Thiessen, V. (2003) *The digital divide in Canadian schools: Factors affecting student access to and use of information technology*. Statistics Canada (Catalogue number 81-597-XIE).

Maki, R., Maki, W., Patterson, M., & Whittaker, D. (2000). Evaluation of a Web-based introductory psychology course: I: Learning and satisfaction in online versus lecture courses. *Behavior research methods, instruments, & computers, 32*(2), 230–239.

Marttunen, M., & Laurinen, L. (2001). Learning of argumentation skills in networked and face-to-face environments. *Journal of Instructional Science, 29*, 127–153.

Mason, B. J., Patry, M., & Bernstein, D. J. (2001). An examination of the equivalence between non-adaptive computer-based and traditional testing. *Journal of Educational Computing Research, 24*(1), 29–39.

Neuhauser, C. (2002). Learning style and effectiveness of online and face-to-face instruction. *The American Journal of Distance Education, 16*(2), 99–113.

Pérez-Prado, A., & Thirunarayanan, M. (2002). A qualitative comparison of online and classroom-based sections of a course: exploring student perspectives. *Educational Media International, 39*(2), 195–202.

Powell, J. V., Aeby, V. G. J., & Carpenter-Aeby, T. (2003). A comparison of student outcomes with and without teacher-facilitated computer-based instruction. *Computers & Education, 40*(2), 183–191.

Ravitz, J., Mergendoller, J., & Rush, W. *What's school got to do with it? Cautionary tales about correlations between student computer use and academic achievement.* Paper presented at AERA 2003, Chicago.

Sankaran, S. R., & Bui, T. X. (2001). Impact of learning strategies and motivation on performance: A study in Web-based instruction. *Journal of Instructional Psychology, 28*(3), 191–203.

Sankaran, S. R., Sankaran, D., & Bui, T. X. (2000). Effect of student attitude to course format on learning performance: An empirical study in Web vs. lecture instruction. *Journal of Instructional Psychology, 27*(1), 66–73.

Schimmoeller, M. Examining engagement, achievement, and interest in science using a personal response computer system. Paper presented at AERA 2003, Chicago.

Schoenfeld-Tacher, R., McConnell, S., & Graham, M. (2001). Do no harm: A comparison of the effects of online vs. traditional delivery media on a science course. *Journal of Science Education and Technology, 10*(3), 257–265.

Sumner, M., & Hostetler, D. (2002). A comparative study of computer conferencing and face-to-face communications in systems designs. *Journal of Interactive Learning Research, 13*(3), 277–291.

Thirunarayanan, M., & Pérez-Prado, A. (2002). Comparing Web-based and classroom-based learning: A quantitative study. *Journal of Research on Technology in Education, 34*(2), 131–137.

Tuckman, B. W. (2002). Evaluating ADAPT: A hybrid instructional model combining Web-based and classroom components. *Computers & Education, 39*(3), 261–269.

Wang, A. Y., & Newlin, M. H. (2000). Characteristics of students who enrol and succeed in psychology Web-based classes. *Journal of Educational Psychology, 92*(1), 137–143.

Wang, A. Y., & Newlin, M. H. (2002). Predictors of Web-student performance: The role of self-efficacy and reasons for taking an online class. *Computers in Human Behavior, 18*(2).

Waschull, S. (2001). The online delivery of psychology courses: Attrition, performance, and evaluation. *Computers in Teaching, 28*(2), 143–147.

Wolverton, M. & Wolverton, M. L. Examining the effectiveness of computer-aided asynchronous augmentation in traditional course delivery. Paper presented at AERA 2003, Chicago.

Yu, F., & Yu, H. (2001). Incorporating e-mail into the learning process: Its impact on student academic achievement and attitudes. *Computers & Education, 38*(1–3), 117–126.

Additional Resources

Allen, G. D. (2001). Online calculus: The course and survey results. *Computers in the Schools, 17*(1), 17–30.

Anastasiades, P. S. (2003). Distance learning in elementary schools in Cyprus: The evaluation methodology and results. *Computers & Education, 40*(1), 17–40.

Annison, J. (2002). Action research: Reviewing the implementation of a distance-learning degree programme utilizing communication and information technologies. *Innovations in Education and Teaching International, 39*(2), 95–106

Blanchette, J. (2001). Questions in the online learning environment. *Journal of Distance Education, 16*(2), 37–57.

Brennan, L. L. (2000). Let your fingers do the talking: Conduction class with "chat." *Journal of Educational Technology Systems, 29*(1), 21–30.

Bryson, M., Petrina, S., Braundy, M., & de Castell, S. (2003). Conditions for success? Gender in technology-intensive courses in British Columbia secondary schools. *Canadian Journal of Science, Mathematics and Technology Education, 3*(2), 185–193.

Chang, C. Y. (2002). Does computer-assisted instruction + problem solving = improved science outcomes? A pioneer study. *The Journal of Educational Research, 95*(3), 143–150.

Chang, K., Sung, Y., & Lee, C. (2003). Web-based collaborative inquiry learning. *Journal of Computer Assisted Learning, 19*, 56–69.

Cifuentes, L., & Shih, Y. (2001). Teaching and learning online: A collaborative between U.S. and Taiwanese students. *Journal of Research on Computing in Education, 33*(4).

Craig, D. V. (2001). View from an electronic learning environment: Perceptions and patterns among students in an online graduate education course. *Journal of Educational Technology Systems, 30*(2), 197–219.

Corbett, B. (2002). Information and communication technology: Access and use. *Education Quarterly Review*, 8–15.

Darkwa, O., & Mazibuko, F. (2002). Creating virtual learning communities in Africa: Challenges and prospects. *First Monday, 5*(5) 12–17.

Dedic, H. (2001). "Do I really hafta?" WebCal, a look at the use of LiveMath Software in Web-based materials that provide interactive engagement in a collaborative learning environment for differential calculus. *Educational Research and Evaluation, 7*(2-3), 285–312.

Fisher, M., & Coleman, B. (2001). Collaborative online learning in virtual discussions. *Journal of Education Technology Systems, 30*(1), 3–17.

Frank, M., Reich, N., & Humphreys, K. (2003). Respecting the human needs of students in the development of e-learning. *Computers & Education, 40*, 57–70.

Gal-Ezer, J., & Lupo, D. (2002). Integrating Internet tools into traditional CS distance education: Students' attitudes. *Computers & Education, 38*(4), 319–329.

Gallini, J. K., & Barron, D. (2001). Participants' perceptions of Web-infused environments: A survey of teaching beliefs, learning approaches, and communication. *Journal of Research on Technology in Education, 34*(2), 139–157.

Gardner, L., Sheridan, D., & White, D. (2002). A Web-based learning and assessment system to support flexible education. *Journal of Computer Assisted Learning, 18*(2), 125–136.

Georgios, D. (2002). *Enhancement of learning process in Web-based courses using combined media components.* [Unpublished Master's Thesis] Naval Postgraduate School, Monterey, CA.

Halang, W. (2002, June 24–27). *Sober thoughts in multimedia based distance learning.* Paper presented at the 24th International Conference on Information technology interfaces, Cavtat, Croatia.

Hargis, J. (2001). Can students learn science using the Internet? *Journal of Research on Computing in Education, 33*(4), 475–487.

Hong, K. S. (2002). Relationships between students' and instructional variables with satisfaction and learning from a Web-based course. *The Internet and Higher Education, 5*, 267–281.

Hughes, M., & Daykin, N. (2002). Toward constructivism: Investigating students' perceptions and learning as a result of using an online environment. *Innovations in Education and Teaching International, 39*(3), 217–224.

Jiang, M. (2000). A study of factors influencing students' perceived learning in a Web-based course environment. *Int. Journal of Educational Telecommunications, 6*(4), 317–338.

Jung, I., Choi, S., Lim, C., & Leem, J. (2002). Effects of different types of interaction on learning achievement, satisfaction and participation in Web-based instruction. *Innovations in Education and Teaching International, 39*(2), 153–162.

Jung, I., & Rha, I. (2000). Effectiveness and cost-effectiveness of online education: A review of the literature. *Educational Technology*, 57–60.

Jung, I., & Rha, I. (2001). A virtual university trial project: Its impact on higher education in South Korea. *Innovations in Education and Teaching International, 38*(1), 31–41.

Kassop, M. (2003). Ten ways online education matches, or surpasses, face-to-face learning. *The Technology Source,* June 10, 2003.

Kim, Y., & Lee, J. (1999). Present and future of cyber education in Korea: Suggestion for its promising future based on representative practices. *International Journal of Education Technology, 1*(1), 51–71.

King, F. B. (2001). Perceptions of technology: A factor in distance education course achievement? *Journal of Educational Computing Research, 24*(4), 407–418.

King, K. P. (2001). Educators revitalize the classroom "bulletin board": A case study of the influence of online dialogue on face-to-face classes from an adult learning perspective. *Journal of Research on Computing in Education, 33*(4).

Land, S. M., & Dornisch, M. M. (2002). A case study of student use of asynchronous bulletin board systems (BBS) to support reflection and evaluation. *Journal of Educational Technology Systems, 20*(4), 265–277.

Lee, M. G. (2001). Profiling students' adaptation styles in Web-based learning. *Computers & Education, 36*, 121–132.

Leh, A. S. (2002). Action research on hybrid courses and their online communities. *Educational Media International, 39*(1), 31–35.

Leonard, J., & Guha, S. (2001). Education at the crossroads: Online teaching and students' perspectives on distance learning. *Journal of Research on Technology in Education, 34*(1), 51–57.

Levin, S., Levin, J., Buell, J., & Waddoups, G. (2002). Curriculum, technology, and education reform (CETER) online: Evaluation of an online master of education program. *TechTrends, 46*(3), 30–38.

Li, Q. (2002). Exploration of collaborative learning and communication in an educational environment using computer-mediated communication. *Journal of Research on Technology in Education, 34*(4), 503–517.

Lin, B., & Hsieh, C. (2001). Web-based teaching and learner control: A research review. *Computers & Education, 37*, 377–386.

Lu, A. (2000). The use and effects of Web-based instruction: Evidence from a single source study. *Journal of Interactive Learning Research, 11*(2), 197–218.

Lupo, D., Erlich, Z. (2001). Computer literacy and applications via distance e-learning. *Computers & Education, 36*(4), 333–345.

Mann, D. (2002). The productivity of learning technologies: A school and learning policy review. *Technos, 11*(1), 12–17.

Mistler-Jackson, M., & Songer, N. B. (2000). Student motivation and Internet technology: Are students empowered to learn science? *Journal of Research in Science Teaching, 37*(5), 459–479.

Monteith, M., & Smith, H. (2001). Learning in a virtual campus: The pedagogical implications of students' experiences. *Innovations in Education and Teaching International, 38*(2), 119–132.

Moor, J., & Zazkis, R. (2000). Learning mathematics in a virtual classroom: Reflection on experiments. *Journal of Computers in Mathematics and Science Teaching, 19*(2), 89–113.

Moore, J. L., Laffey, J. M., Espinosa, L., & Lodree, A. W. (2002). Bridging the digital divide for at-risk students: Lessons learned. *TechTrends, 46*(2), 5–9.

Murfin, B. (2001). A case study of math and science teacher education in a collaborative virtual learning environment. *Journal of Computers in Mathematics and Science Teaching, 20*(4), 405–425.

Niederhauser, D. S., Reynolds, R. E., Salmen, D. J., & Skolmoski, P. (2000). The influence of cognitive load on learning from hypertext. *Journal of Educational Computing Research, 23*(3), 237–255.

Oliver, R., & Omari, A. (2001). Student responses to collaborating and learning in a Web-based environment. *Journal of Computer Assisted Learning, 17*, 34–47.

Plass, J. L., & Whelan, R. (2002). Research on Web-based learning and instruction: Review of literature from 1993 to 2001. Unpublished manuscript.

Pollacia, L. F., & Simpson, C. (2000). Web-based delivery of information technology courses. *Journal of Education Technology Systems, 29*(1), 31–40.

Reynolds, K. M. (2002). The effectiveness of a listserv for teacher preparation. *Journal of Educational Technology Systems, 31*(1), 71–87.

Robbins, N., Lo, Y., Hou, F., Chou, T., Chen, C., Chen, C., Chen, W., Chen, Y., Wang, S., Huang, S., & Lii, J. (2002). Computer-mediated communication and University International students. *Journal of Educational Technology Systems, 30*(3), 323–331.

Sakamoto, T. (2002). E-learning and educational innovation in higher education in Japan. *Educational Media International, 39*(1), 9–16.

Sanders, D. W., & Morrison-Shetlar, A. I. (2001). Student attitudes toward Web-enhanced instruction in an introductory biology course. *Journal of Research on Computing in Education, 33*(3), 251–262.

Schoech, D. (2000). Teaching over the Internet: Results of one doctoral course. *Research on Social Work Practice, 10*(4), 467–486.

Seng, L., & Mohamad, F. S. (2002). On-line learning, Is it meant for science courses? *The Internet and Higher Education, 5*, 109–118.

Shedletsky, L. J., & Aitken, J. E. (2001). The paradoxes of online academic work. *Communication Education, 50*(3), 206–217.

Shuell, T. J., & Farber, S. L. (2001). Students' perceptions of technology use in college courses. *Journal of Educational Computing Research, 24*(2), 119–138.

Smith, G. G., Ferguson, D., & Caris, M. (2002). Teaching online versus face-to-face. *Journal of Educational Technology Systems, 30*(4), 337–364.

Soong, M., Chan, H., Chua, B., & Loh, K. (2001). Critical success factors for online course resources. *Computers & Education, 36*(2), 101–120.

Stocks, J. T., & Freddolino, P. P. (2000). Enhancing computer-mediated teaching through interactivity: The second iteration of a World Wide Web-based graduate social work course. *Research on Social Work Practice, 10*(4), 505–518.

Swan, K. (2001). Virtual interaction: Design factors affecting student satisfaction and perceived learning in asynchronous online courses. *Distance Education, 22*(2), 306–331.

Tansley, C., & Bryson, C. (2002). Virtual seminars: A viable substitute for traditional approaches? *Innovations in Education and Training International, 37*(4), 335–345.

Thomas, M. J. W. (2002). Learning within incoherent structures: The space of online discussion forums. *Journal of Computer Assisted Learning, 18*(3), 351–366.

Tiene, D. (2000). Online discussions: A survey of advantages and disadvantages compared to face-to-face discussions. *Journal of Educational Multimedia and Hypermedia, 9*(4), 369–382.

Verbeeten, M. J. (2001). Learner-centered? It's just a click away... *Journal of Educational Technology Systems, 30*(3), 159–170.

Volery, T. (2001). Online education: An exploratory study into success factors. *Journal of Educational Computing Research, 24*(1), 77–92.

Vonderwell, S. (2003). An examination of asynchronous communication experiences and perspectives of students in an online course: A case study. *The Internet and Higher Education, 6*, 77–90.

Warschauer, M., & Meskill, C. (2000). Technology and second language learning. In J. Rosenthal (Ed.), *Handbook of undergraduate second language education* (pp. 303–318). Mahwah, NJ: Lawrence Erlbaum.

Woodrow, J. E. J., Mayer-Smith, J. A., & Pedretti, E. G. (2000). Assessing technology-enhanced instruction: A case study in secondary science. *Journal of Educational Computing Research, 23*(1), 15–39.

Yang, S. C. (2001). Integrating computer-mediated tools into the language curriculum. *Journal of Computer Assisted Learning, 17*, 85–93.

Yang, S. C. (2001). Language learning on the World Wide Web: An investigation of EFL learners' attitudes and perceptions. *Journal of Educational Computing Research, 24*(2), 155–181.

Appendix A: Contacts

Arafeh, Sousan	USA
Bialo, Ellen	USA
Brown, Jonathan	Canada
Bryson, Mary	Canada
Cano, Edwin A. Alvarado	Guatemala
Carey, Tom	Canada
Carty, Tony	Canada
Cheng, Chee Hing	Hong Kong
Claro, Magdalena	Chile
Cuneo, Carl	Canada
Escorcia, German	
Gardner, John	Ireland
Goddard, Ron	Canada
Gumley, Gary	Canada
Hasan, Abrar	France
Karp, Naomi	USA
Kennedy, Robert	Canada
Kim, Doo Jung	South Korea
LaForest, Lynda	Canada
Longworth, Richard	Canada
Mann, Dale	USA
Mutzig, Jean-Marc	Brazil
Nordwall, Margareta	Sweden
Ogawa, Yo	Japan
Pont, Beatriz	France
Ramjewon, Toolseram	Mauritius
Sobey, Morton	Norway
Sweet, Richard	France
Tam, Peter	Hong Kong
Varela, Carmen	Paraquay
Warschauer, Mark	USA
Willms, Doug	Canada

Appendix B: Journals

American Educational Research Journal
American Journal of Distance Education
British Journal of Educational Technology
Cambridge Journal of Education
Child Development
Cognition and Instruction
Communication Education
Computers & Education
Computers in Human Behavior
Computers in Human Services
Computers in the Schools
Computing Research
Distance Education
Distance Learning
Education Canada
Educational Media International
Education Quarterly Review
Educational Computing and Technology
Educational Computing Research
Educational Leadership
Educational Psychology
Educational Research
Educational Research and Evaluation
Educational Researcher
Educational Technology
Educational Technology Research and Development
Educational Technology Review
European Journal of Psychology of Education
Group Decision and Negotiation
Information Society
Innovations in Education and Teaching International
Instructional Science
Interactive Learning Environments
International Journal of Educational Technology
International Journal of Educational Telecommunications

International Journal of Human-Computer Studies
International Journal of Instructional Media
International Journal on E-Learning
Internet and Higher Education
Interpersonal Computing and Technology
Journal of Applied Developmental Psychology
Journal of Asynchronous Learning Networks
Journal of Computer Assisted Learning
Journal of Computer Based Instruction
Journal of Computers in Mathematics and Science Teaching
Journal of Computing in Childhood Education
Journal of Distance Education
Journal of Economic Education
Journal of Educational Computing Research
Journal of Educational Media
Journal of Educational Multimedia and Hypermedia
Journal of Educational Psychology
Journal of Educational Research
Journal of Educational Technology Systems
Journal of Experimental Child Psychology
Journal of Experimental Education
Journal of Human Behaviour and Learning
Journal of Information Technology for Teacher Education
Journal of Instructional Psychology
Journal of Interactive Learning Research
Journal of Interactive Media in Education
Journal of Research in Childhood Education
Journal of Research in Reading
Journal of Research in Science Teaching
Journal of Research on Computing in Education
Journal of Research on Technology in Education
Journal of School Psychology
Journal of Technology in Human Services
Journal of the American Society for Information Science
Journal of Visual Literacy
Journal of Universal Computer Science
Learning and Instruction
Learning Environments Research
New Directions for Adult and Continuing Education

New Directions for Teaching and Learning
Personal Technologies Journal
Quarterly Review of Distance Education
Research in Science and Technological Education
Research on Social Work Practice (2000)
Review of Educational Research
Science Communication
Science Education
Simulation and Gaming
Teaching of Psychology
Tech Trends
Technology
Voprosy-Psychologii

Appendix C: List of Data Elements for Coding

Coding Scheme
Author and title:
Identification of report:
- Citation
- Contact
- Hand search
- Electronic database
- Other

Source peer reviewed? Yes/No

Country of Study:

Setting:
- Elementary
- Secondary
- University
- Workplace
- Other

Subjects
- Gender
  - Female
  - Male
- Age:
  - Elementary
  - Secondary
  - University (adults)

Method

- Type
  - Qualitative
    - Ethnographic
    - Survey
    - Case study
    - Other
  - Quantitative
    - Survey
    - RCT
    - Quasi-experimental
    - Other

Control/Comparison Group
- Yes
- No

Random Assignment
- Yes
- No

Blind Subjects
- Yes
- No

Blind Investigator
- Yes
- No

Technologies
- Internet
- Networked
- Tele-teaching
- E-mail
- CMC

Results

Effect size
- Yes
- No
- Calculable?