

**A NATURALISTIC JOURNEY INTO THE COLLABORATORY: IN SEARCH
OF UNDERSTANDING FOR PROSPECTIVE PARTICIPANTS**

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

IN THE GRADUATE SCHOOL OF THE

TEXAS WOMAN'S UNIVERSITY

SCHOOL OF LIBRARY AND INFORMATION STUDIES

BY

joanne twining, A.S., B.S., M.L.S.

DENTON, TEXAS

DECEMBER 1999

TEXAS WOMAN'S UNIVERSITY
DENTON, TEXAS

November 19, 1999

To the Associate Vice President for Research and Dean of the Graduate School:

I am submitting herewith a dissertation written by joanne twining entitled "**A Naturalistic Journey into the Collaboratory: In Search of Understanding for Prospective Participants.**" I have examined this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Library and Information Studies.

John D'Angelo, Ph.D., Major Professor

We have read this dissertation and recommend its acceptance:

Joy McGregor, Ph.D.

Don Edwards, Ph.D.

Keith Swigger, Ph.D., Dean of the School

Accepted

Associate Vice President for Research
and Dean of the Graduate School

Copyright © joanne twining, 2000
All Rights Reserved

ACKNOWLEDGEMENTS

No one journeys alone. Thank you, Joe Natale, for digging deep and persistently, as all good librarians should. Thank you, Jason O'Brien, my son and best friend, for the inspiration, encouragement, emotional and intellectual support, for the many hours of delicious brain-picking, and for your extraordinary insights and brilliance, without which this work would have been something altogether different, and not nearly so much fun☺

Thank you, Geraldine Rinna and Dianne Michael Frank for being unabashedly proud and providing unlimited emotional support through the personal challenges simultaneous with this work. I treasure your counsel and envy your remarkable intelligence and harmony, dear sisters.

Thank you, Dr. Keith Swigger, Dean of "my" Library School, trusted mentor and gadfly, for the promise delivered. With but a single word, you often rocked my reality, and always, with gentle endurance, guided me through the most challenging adventure of my life: coming to the edge of my knowledge and facing the abyss. The world is a much better place for the librarians you have brought to it, and I will owe you forever for the privilege of being among them.

Thank you committee members, Dr. Don Edwards, my numbered, wired one, and Dr. Joy McGregor, my intersubjective sensitive, for your quiet support, your keen eyes, your extraordinary insights, your patient arguments, and for being exquisite examples of the American scholar. Would that I achieve a fraction of your wisdom and grace.

Thank you, Dr. John D'Angelo, Dr. D., my Dissertation Chair and most valued advisor. You are an outstanding teacher, a sensitive mentor, a superlative guide, and the consummate collaborator. Your confidence allowed me the extraordinary privilege of exploring my mind's potential, and you have taught me well, and better than anyone else, ever. You rock.

And finally, thank you Rachael Green and Kevin Miles, fellow students, for letting me talk, talk, talk what must have seemed gibberish as this work was taking form.

I could have done it without any of you, but it wouldn't have been the same.

ABSTRACT

A Naturalistic Inquiry Into The Collaboratory: In Search Of Understanding For Prospective Participants

joanne twining
December 1999

This study is a three-phase naturalistic probe of the information environment of the collaboratory in support of the expansive, ecological research of others. The collaboratory is a "center without walls," or a virtually collocated, collaborative laboratory where scientists, instruments, and data come together via computer network technology without regard to geographic location. In Phase One, an objective reality of the collaboratory is constructed from the documents made accessible through the world's libraries. Taxonomy construction and quantitative and qualitative data analysis investigate and prove as practiced principles the assumptions of relative equality of contribution to collaboratory research by the hard and soft sciences, and the inherent interdisciplinarity of the collaboratory environment. An emergent theory of the collaboratory as an ungendered environment is developed. Phase Two creates a subjective reality of the collaboratory based on experiential immersion in the online environment. An evaluative instrument, the CIRAL matrix of criteria for inclusion as a collaboratory, is developed and tested, and four collaboratory site visits are

developed. The collaboratory is found to be an instrumentally-determined social environment, with each implementation unique in its combinations of communication modes and media, and each generating individualized data stores. Phase Three constructs an intersubjective reality of the collaboratory during an electronic Delphi among collaboratory pioneers. The Delphi determines the "rules of the road" for the collaboratory and identifies skills collaboratory pioneers value in prospective participants. Phase Three identifies cognitive dissonance between the intersubjective reality of collaboratory pioneers and Phase One's findings of relative equality of contribution to, and ungenderedness of the collaboratory environment. Size of collaboratory is explored as a determining factor in preferences for balance between formal and informal communication modes, and structured and fluid experiment planning. The three realities are intertwined to construct a holistic, synoptic survey of the collaboratory as an emergent knowledge environment in which old science is done with new tools, but from which new science has yet emerged.

CONTENTS

ACKNOWLEDGEMENTS	iv
ABSTRACT	vi
ILLUSTRATIONS	xi
LIST OF TABLES	xii
CHAPTER ONE	1
Research Agenda	2
Statement of the Problem	8
Paradigm of the Study	9
Methodology	12
Phase One Toward an Objective Reality of the Collaboratory	13
CHAPTER TWO Foundation Documents	14
The Collaboratory Literature	15
Philosophical and Intellectual Foundations	18
The National Collaboratory -- White Paper (Wulf 1988).....	19
Towards a National Collaboratory (Lederberg and Uncapher 1989).....	26
Conclusion	39
CHAPTER THREE Discipline, Focus, and Type of Publications	41
The Document Retrieval Set	41
Taxonomy Construction	45
The Wulf Taxonomy and Analysis	46
The Haddow Taxonomy and Analysis	52
Descriptive Statistical Analysis	57
Discipline x Focus x Article Type Co-analysis.....	58
Conclusion	65
CHAPTER FOUR Topic and Approach of Publications	68
The Lederberg and Uncapher Taxonomy	69
Lederberg and Uncapher's "Cycles"	75
Theory-type Research Articles by Discipline.....	77
Conclusion	81
CHAPTER FIVE Toward a Theory of the Collaboratory	83
Systems Architecture Theory-Type Research	84
Tools and Technology Theory-Type Research	85
Uses and Testbed Theory-Type Research	87
Theory Development.....	90

Conclusions	96
CONCLUSION OF PHASE ONE.....	98
Phase Two Toward A Subjective Reality Of The Collaboratory Error!	
Bookmark not defined.	
CHAPTER SIX Criteria for Inclusion as Collaboratory	Error! Bookmark not defined.
defined.	
National Collaboratories (NRC 1993)	Error! Bookmark not defined.
The Generic Model of the Collaboratory	Error! Bookmark not defined.
The CIRAL Matrix	Error! Bookmark not defined.
Discussion of Criteria for Inclusion.....	Error! Bookmark not defined.
Conclusion.....	Error! Bookmark not defined.
CHAPTER SEVEN Into the Collaboratory.....	Error! Bookmark not defined.
M2C, The Materials MicroCharacterization Collaboratory	Error! Bookmark not defined.
not defined.	
TPM : TelePresence Microscopy	Error! Bookmark not defined.
NIST TelePresence Microscopy Site	Error! Bookmark not defined.
Conclusion.....	Error! Bookmark not defined.
CHAPTER EIGHT Descriptive Studies.....	Error! Bookmark not defined.
Derivative "Collaboratories"	Error! Bookmark not defined.
SPARC : The Space Physics and Aeronomy Research Collaboratory	Error! Bookmark not defined.
Bookmark not defined.	
DOE2000	Error! Bookmark not defined.
REE Collaboratory Testbed at the DIII-D Tokamak	Error! Bookmark not defined.
defined.	
EMSL : The Environmental Molecular Science Collaboratory	Error!
Bookmark not defined.	
Conclusion.....	Error! Bookmark not defined.
CONCLUSION OF PHASE TWO.....	Error! Bookmark not defined.
Phase Three Toward an Intersubjective Reality of the Collaboratory Error!	
Bookmark not defined.	
CHAPTER NINE Electronic Delphi Among Collaboratory Pioneers	Error!
Bookmark not defined.	
"Rules of the Road"	Error! Bookmark not defined.
The Delphi Method	Error! Bookmark not defined.
Collaboratory Pioneers	Error! Bookmark not defined.
Round One	Error! Bookmark not defined.

Round Two	Error! Bookmark not defined.
Conclusion of Phase Three.....	Error! Bookmark not defined.
CONCLUSION OF THE STUDY	Error! Bookmark not defined.
Research Needs	Error! Bookmark not defined.
REFERENCES.....	Error! Bookmark not defined.
APPENDIX A The Collaboratory Literature Retrieval Set, Annotated	Error!
Bookmark not defined.	
APPENDIX B The CIRAL Matrix for Inclusion as a Collaboratory	Error!
Bookmark not defined.	
APPENDIX C Phase Two Consent Form	Error! Bookmark not defined.
APPENDIX D Phase Three Consent Form	Error! Bookmark not defined.
APPENDIX E Round One Delphi Questionnaire	Error! Bookmark not defined.
APPENDIX F Round Two Delphi Instrument.....	Error! Bookmark not defined.

ILLUSTRATIONS

Figure 1. Publications by Year	44
Figure 2. Percentage of Publications by Discipline	50
Figure 3. Publications by Discipline and Focus.....	51
Figure 4. Article Type.....	54
Figure 5. Article Type and Year	54
Figure 6. Trends Analysis, Publications over Time	57
Figure 7. Focus of Research Articles.....	59
Figure 8. Discipline of Research Articles	60
Figure 9. Percent of Total and Research Publications by Discipline	63
Figure 10. Percentage of Total and Research-Type Articles, by Discipline	64
Figure 11. Number of Total and Research-Type Articles, by Discipline	65
Figure 12. Topic of Articles	73
Figure 13. Approach of Articles.....	74
Figure 14. Topic and Approach of Articles.....	74
Figure 15. Approach of Articles Annually	76
Figure 16. Theory-Type Research Articles by Topic.....	77
Figure 17. Theory -Type Research by Topic and Discipline	78
Figure 18. Theory-Type Research Articles by Topic and Combined Disciplines.	81
Figure 19. M2C Collaboratory Logo	Error! Bookmark not defined.
Figure 20. Screenshot of M2C TelePresence Microscopy Session	Error!
Bookmark not defined.	
Figure 21. The Sondrestrom Facility	Error! Bookmark not defined.
Figure 22. SPARC Website	Error! Bookmark not defined.
Figure 23. Inside the DIII-D Tokamak	Error! Bookmark not defined.
Figure 24. EMSL Logo: Borromean Rings	Error! Bookmark not defined.
Figure 25. Communication Modes of Collaboratory Types	Error! Bookmark not defined.
defined.	
Figure 26. EMSL CURE Session	Error! Bookmark not defined.

LIST OF TABLES

Table 1. Basic Taxonomy.....	45
Table 2. Taxonomy of Disciplines	48
Table 3. Publications by Discipline and Focus.....	49
Table 4. Publications by Type.....	53
Table 5. Frequencies and Percentages of Articles by Discipline	61
Table 6. Topic and Approach of Articles	71
Table 7. M2C Collaboratory Criteria for Inclusion .	Error! Bookmark not defined.
Table 8. Comparison of Collaboratory Data Types	Error! Bookmark not defined.

CHAPTER ONE

This study explores the information environment of collaboratory from three perspectives --the objective, the subjective, and the intersubjective--to create a synoptic survey in support of expansive studies of the collaboratory's larger information ecology. Information ecologies are "responsible, informed, engaged interactions among people and advanced information technologies" (Nardi and O'Day 1999, 24). The key constituents of an information ecology are people, practices, values, and technology (Nardi and O'Day 1999, 60). An information environment is the aggregate of surrounding things, conditions, or influences. This study assumes the collaboratory is an information ecology, and therefore has a discernable information environment.

The collaboratory is a "center without walls," an online, electronic environment where scientists, instruments, and data come together via computerized networks (NRC 1993); where production and dissemination of scientific discovery will be quickened (Wulf 1988); and where the capabilities of the human intellect will be amplified (Lederberg and Uncapher 1989). The collaboratory has been a part of the information research agenda of the United States for a decade: since 1988 when the word "collaboratory" was first coined. The Internet, or global network of networks, is, in part, a product of collaboratory-

funded research. While the collaboratory uses Internet technologies, it is not the Internet, and the Internet is not it. Both are part of the larger National Information Infrastructure (NII) and the Global Information Infrastructure (GII)., and while all digital initiatives are intertwined, the collaboratory is a unique and distinct ecology with a unique and distinct environment. To date, no general environmental survey of the collaboratory has been undertaken. This study addresses that need.

Research Agenda

This study takes three distinct approaches to the collaboratory environment. In Phase One, an objective reality is constructed based on the documents made available through the world's libraries. In Phase Two, a subjective reality is constructed during immersion in the online environment of the collaboratory. In Phase Three, an intersubjective reality is constructed based on interaction with Collaboratory Pioneers. Together, the objective, subjective, and intersubjective approaches construct a general environmental overview of the collaboratory.

Phase One creates an objective reality by consulting the only enduring public record of the collaboratory's first decade: the publications held by the world's libraries. Two taxonomies, or quantitative classification schemes, based on collaboratory foundation documents (Wulf 1988, Lederberg and Uncapher 1989),

and a third based on Haddow's (1997) types of publications, are constructed. The collaboratory literature (n=86) is analyzed and cross-analyzed by discipline, focus, type, topic, and approach of research and publications. The assumptions of relative equality of contribution to research leading to development of the collaboratory and the interdisciplinarity the collaboratory environment are explored and proved as principles by practices reflected in the literature. Phase One concludes with a synoptic analysis of a subset of Theory-Type Research publications (n=22) from which a theory of the collaboratory as an ungendered environment emerges.

Phase Two creates a subjective, experiential reality during prolonged immersion in the online environment of the collaboratory. Criteria for inclusion as a collaboratory are developed and aspects and elements of four collaboratory environments are explored and discussed. Phase Three creates an intersubjective reality of the collaboratory during an electronic Delphi among collaboratory pioneers to determine the "rules of the road" for the collaboratory and identify skills collaboratory pioneers value in prospective participants.

This environmental survey serves the progress of collaboratory science by creating a resource to aid general understanding of prospective participants. Understanding the concept of the collaboratory is important to the advance of virtual science generally, and so is of interest to higher education, libraries,

business, the professions, as well as to scientists and individuals who may come to participate in this new knowledge environment.

The once distinctly sequential document-based, experience-based, and colleague-based information environments converge in the collaboratory. This convergence has logical consequences for the process of collaboratory science beyond the claim that constraints of time and space will be overcome. New information environments may require new knowledge processes. Examining the environment should shed light on those processes. The relative newness of the collaboratory provides a unique opportunity to explore the convergence using the traditionally sequential objective, subjective, and intersubjective approaches.

Phase One's objective reality determines how the collaboratory is represented in the library literature. Knowledge created in Phase One should reflect what any researcher consulting the cultural record might come to know and understand about the collaboratory's first decade. Phase Two determines if the collaboratory exists as an environment as it is represented in the library literature. Knowledge created in Phase Two should reflect what any researcher experiencing the online collaboratory might come to understand. Phase Three determines the "rules of the road" for the collaboratory and identifies skills collaboratory pioneers value in prospective participants. The "rules of the road" are a previously unaddressed information need identified in the National Research Council's (1993) *National*

Collaboratories: Applying Information Technologies for Scientific Research.

Knowledge created in Phase Three reflects what any researcher engaged with collaboratory participants might come to know and understand about working in the collaboratory.

Isolating the three distinct approaches is significant for several reasons. Understanding is a human process by which information is transformed into knowledge (Nitecki 1993). There are many philosophies to guide how knowledge is constructed (Crimshaw 1986, Wilson 1999). Traditionally, scholarly research is achieved in a three step process: by studying the literature, as in library research; by praxis, as in the laboratory experience; and by interaction and exchange with individuals and groups, as at conferences or in the scholarly peer review process. Granted, these processes, by necessity, overlap in the traditional scholarly environment. How knowledge is created in an environment purposively constructed to simultaneously intertwine the objective, subjective, and intersubjective approaches should shed light on the new processes the collaboratory intends to facilitate.

A triangulated approach (Nitecki 1995) provides a logically balanced intellectual alternative to the more traditional dichotomous and disproving approach of the scientific method (Wilson 1999). The notion of triangulated

thinking is reflected in the "see one, do one, teach one" ideal of American education and in the older Chinese wisdom:

Tell me and I forget.
Show me and I remember.
Involve me and I understand. (Lecture 1995)

As access to the collaboratory becomes more widely available, this study serves by providing a general introduction for prospective participants. A synoptic, environmental exploration will be useful on other fronts, as well. Historians, toolmakers, and other researchers may use this study to support inquiry into other aspects and elements of the collaboratory's larger information ecology. Librarians and information professionals may find this study useful as they enter the collaboratory environment to work. Business, economics, publishing, and the social sciences may find it useful as the concept of virtual work spreads to their less instrument-dependent environments and they become part of the "intellectual stew" the collaboratory intends to produce.

The cycle of traditional scholarly publishing produces a significant delay between discovery, understanding, and dissemination of findings. The collaboratory is often represented as a way to speed up, if not side step, portions of this process and hasten the research-to-shrink-wrap cycle. The delay in scholarly publishing is regularly a year, and frequently several years, by which time, particularly in the fast-paced world of technological development,

publication may serve primarily as historical background rather than to practically inform ongoing research of scientists and scholars (Schrage 1991). No general investigation of the collaboratory's alleged ability to sustain and facilitate productive exchange and co-manipulation of active datasets while facilitating mutual discovery, and continuous real-time problem solving, exists. This study is a first step in that inquiry.

To an item, every scholarly account of the collaboratory is from an insider's perspective; from someone involved and vested in some aspect of the collaboratory's creation or development. There is no scholarly account of a journey into the collaboratory by someone from outside its walls. This study provides that perspective.

Phase One seeks to understand what the collaboratory says it is, and so provides a foundation from which the published account and the actual experience of the collaboratory may be evaluated. In Phase One, three taxonomies are constructed and the collaboratory literature is classified, then analyzed within and among these taxonomies. This taxono-bibliometric analysis represents a type of tacit collaboration of minds across time reached by searching, reading, and analyzing the published literature. It counts no experience of the environment, nor does it count on the bilateral exchange of

ideas, or formal collaborations over time, all three of which the collaboratory, by design, intends to simultaneously foster.

Phase Two seeks to understand the technologically-enabled environment of the collaboratory, and to determine whether and to what extent the collaboratory vision has been achieved. Phase Three seeks to understand the personality and emergent culture of the collaboratory through purposive interaction with collaboratory pioneers.

Statement of the Problem

The larger agenda of this research is understanding, and the motivation is curiosity. What is the collaboratory? Does it exist? What are its elements and attributes? What is its nature? What does the collaboratory do? What can you do in a collaboratory? What is it like to participate in collaboratory activities? What types of research are being conducted in the collaboratory? For what other activities might collaboratory technologies be used? Will the collaboratory indeed change the way science is conducted? How might it change the way science is conducted? Why do scientists collaborate online?

The problem delegated to Phase One of this study is: How does the library portray the collaboratory? Does that portrayal reflect the ideals put forth in the

collaboratory's foundation documents? What are the collaboratory's theoretical foundations? What does the published record of collaboratory research reveal?

The problem delegated to Phase Two of this study is: Is the collaboratory as the library portrays? Are the ideals and philosophies put forth in the literature reflected in the actual environment?

The problem delegated to Phase Three of this study is: What do collaboratory pioneers say are the "rules of the road" for the collaboratory? What skills do collaboratory pioneers value in prospective participants? What does the bilateral exchange of ideas about the collaboratory with collaboratory pioneers reveal?

Taken as a whole, these perspectives should provide a general, environmental overview of the collaboratory. The process of probing the environment in search of understanding should shed light on how knowledge might be created as the objective, subjective, and intersubjective fuse online.

Paradigm of the Study

This study employs the naturalistic/constructivist paradigm (Erlandson, et al. 1993), and rests firmly on the notion that

shared constructions, developed collaboratively by empowered individuals, are the basis for significant cross-cultural and interpersonal understandings. (xvii)

The naturalistic/constructivist paradigm assumes no single reality exists and admits ungeneralizable, context-specific subjectivity as an appropriate process of inquiry. The aim of the naturalistic paradigm is to illuminate a single, specific context and provide assumptions, principles, working hypotheses and emergent theory for the expansive research of others. Naturalistic inquiry allows methods to evolve during the course of research rather than requiring that methodologies be determined ahead of the research. Naturalistic inquiry must meet the criteria of "trustworthiness" (Erlandson, et al 1993) as defined by credibility, transferability, dependability, and confirmability (Guba 1981, Guba and Lincoln 1981, Guba and Lincoln 1989, Lincoln and Guba 1985).

Credibility is defined as the "degree of confidence in the 'truth' that the findings of a particular inquiry have for the subject with which--and context within which--the inquiry is carried out" (Erlandson, et al. 1993, 29). Credibility is achieved by strategies including prolonged engagement, persistent observation, triangulation, referential adequacy materials, peer debriefing, and member checks.

Transferability is defined as "the extent to which its findings can be applied in other contexts or with other respondents" (Erlandson, et al. 1993, 31). Transferability and generalizability are not the same thing. Generalizable findings must apply across all environments while transferability allows knowledge gained

to be applied to other environments. Transferability is achieved two ways: through thick description of sufficient detail and precision that it brings the reader vicariously into the environment under investigation, and through purposive sampling governed by emerging insights and information achieved during the course of the investigation.

Dependability is defined as the extent to which, if the inquiry "were replicated with the same or similar respondents (subjects) in the same (or similar) context, its findings would be repeated" (Erlandson, et al. 1993, 33). Dependability is achieved through a "dependability audit" which includes construction and maintenance of an archive facilitating access to all documentation as well as a running account of the process of inquiry in the form of researcher logs.

Confirmability of an inquiry is defined as "the degree to which its findings are the product of the focus of its inquiry and not of the biases of the researcher" (Erlandson, et al. 1993, 34). Confirmability is achieved when constructions, assertions, and facts can be tracked to their original sources and when the logic behind their construction leads to an explicit and implicitly coherent and corroborating whole. Confirmability is achieved via a "confirmability audit" which allows external reviews to judge the conclusions, interpretations and recommendations of the inquiry. The Dependability Audit and the Confirmability Audits are facilitated by the construction of a project library using commonly

available relational database, spreadsheet and word processing software, and Internet technologies.

Methodology

Phase One of this study begins with a comprehensive search of the world's libraries for documents pertinent to the collaboratory, and proceeds without further design, trusting the document retrieval set to reveal how best it might be understood. Therefore, specific methodologies emerge during, and are discussed as part of, the analysis of data. Phase Two chronicles the journey into the online environment of the collaboratory using thick description during prolonged immersion, and is also guided by discoveries made during the course of the research. Phase Three relies on an electronic permutation of the Delphi Technique (Linstone and Turoff 1975) which employs the process of iterative interrogation and reduction to consensus, stasis, or understanding among purposively selected authorities who are most likely able to answer the questions raised.

Phase One

Toward an Objective Reality of the Collaboratory

CHAPTER TWO

Foundation Documents

The philosophical, intellectual and instrumental foundations of the collaboratory are provided by three key documents. Two of these documents are examined in detail in this chapter. They provide the foundation on which Phase One's analysis rests. Three taxonomies, or quantitative classification schemes, are constructed, and the collaboratory literature (n=86) is analyzed and cross analyzed by discipline, focus, type, topic, and approach of research and publications. The assumptions, practices, and principles of the collaboratory, as put forth in its foundation documents and reflected by the library literature, are explored. The content of a subset of Theory-Type Research publications (n=22) is qualitatively analyzed and an emergent theory of the collaboratory as an ungendered, environmental is put forth.

The first key document, *The National Collaboratory – A White Paper* (Wulf 1988) sets the philosophical foundation of the collaboratory and identifies the disciplines that need to contribute to, and the focus of the research needed for development of the collaboratory. The second document, *Towards a National Collaboratory: Report of an Invitational Workshop at the Rockefeller University*

March 13-15, 1989 (Lederberg and Uncapher), provides the intellectual foundation for the collaboratory. It outlines the National Science Foundation's National Collaboratory research agenda, and identifies three topics and approaches of research needed for development of a National Collaboratory. Neither document is published, nor is either available from any lending library. Nevertheless, they are the most widely cited publications in the collaboratory literature.¹

The Collaboratory Literature

For the purpose of this study, the collaboratory literature is defined as resources available through the intermediation of the library. The collaboratory literature includes print and electronic papers and journals, books, reports, and microfilm and microform documents, as well as electronic databases of collected documents, or document surrogates. It does not include private, uncirculated

¹ The Wulf *White Paper* was first identified as cited-in the Lederberg and Uncapher report in a footnote to one of the retrieval set documents. But, neither document was cataloged as held by any lending library in the world, nor were they indexed as available in any database; neither was either available directly from the National Science Foundation, or from the authors. The Lederberg and Uncapher report, which contains the *White Paper* as an appendix, was eventually tracked down by an enterprising access librarian at TWU, Joe Natale, who called upon a librarian at Rockefeller University, who gladly descended into the bowels of the library, located a dusty box of documents remaining from the 1989 workshop, and thumbed through the entire contents until she located the report, which she duplicated and delivered.

documents or correspondences, interactive online environments, or documents published and available without library intermediation, such as commercially available books not held, and documents on the World Wide Web or other parts of the public Internet. Documentary evidence is revealed during interaction with the objects or artifacts of the library during the process of “library research.”

The documents included in the collaboratory literature are those to which access is gained by use of the unadulterated search string “collaboratory.” This study does not use truncated or wildcard derivatives of the word. The use of truncated or wildcard derivatives produces a huge field of documents that are closely related, or relevant to the concept of the collaboratory, but not pertinent to this study. Such documents are, for instance, those dealing with the act of *collaboration*, the *collaborative* attitude, or people who are *collaborators*. While these and other concepts are certainly relevant to the development and use of the collaboratory generally, and will be helpful as this agenda progresses, they open a literature base beyond the intent of this study. This study focuses exclusively and sharply on the collaboratory as an information environment. The boundaries of that environment are negotiated as the study progresses.

There is also substantial preceding, succeeding, simultaneous, and derivative literature with strong relational ties to the collaboratory. Those documents are also not included in this study. Barua (1995) provides an analysis of the near

meteoric development and evolution of the vocabulary, concepts, and technologies leading to the collaboratory. Access to this substantial literature may be achieved using keywords and phrases including Decision Support Systems (DSS), Computer-based Systems for Cooperative Work (CSCW), Group Support Systems (GSS), Group Decision Support Systems (GDSS), GroupWare Systems, Computer Mediated Communication (CMC), and others. Related concepts may also be accessed using the words and phrases DARPA, ARPA, NREN, World Wide Web, Internet, Digital Library, National Information Infrastructure (NII), Global Information Infrastructure (GII), and their associated concepts.

The intent of Phase One is to focus sharply and exclusively on those documents that are highly pertinent to and specifically address the collaboratory as an environment. The documentary evidence of the collaboratory begins in May 1988, when Wulf coined the word, and, for the purpose of Phase One, ends in December 1998. Pertinent documents obtained outside library research, or outside the time frame of Phase One, are addressed, when appropriate, in Phases Two and Three of this study.

Philosophical and Intellectual Foundations

The documented story of the collaboratory began in March 1989 at an invitational workshop convened by Dr. William A. Wulf, then Director of the National Science Foundation's (NSF) Directorate for Computer and Information Science and Engineering (CISE). Wulf gathered twenty-nine scientists and researchers to Rockefeller University in New York, and charged them with developing a research agenda to actualize a National Collaboratory.

Wulf developed the notion of the collaboratory in a December 20, 1988 *White Paper* for the NSF, which he presented to the conference (Wulf 1988, Appendix A in Lederberg and Uncapher 1989). Wulf footnoted in the *White Paper* that the word collaboratory was "invented to combine the words collaboration and laboratory" (2).

Much in the same way that H.G. Wells' (1938) *World Brain* visualized a "central intellectual organism" (Rayward 1999), and Vannevar Bush's (1945) *Atlantic Monthly* article visualized the memex machine and the hyperdocument environment we now recognize as the World Wide Web, Wulf's collaboratory vision promised fundamental changes in the way science is conducted. Wulf incorporated access to and remote manipulation of rare and expensive scientific instruments along with interactive human knowledge networks incorporating real-time and document-based communication of various sorts. The collaboratory is

a work environment that brings geographically dispersed scientists, instruments, and data together in a simultaneously “live” technologically-enabled environment.

The vision of the collaboratory did not wholly spring from Wulf, however (Banks 1993, Robbin 1995). Yet unnamed, the concept spawned almost simultaneously in several quarters of government and military research in the mid-1980s. Computer and information scientists working on the logistics, languages, architectures, and technicalities of what would eventually emerge as the Internet began to imagine potential uses and use-based environments distinct from, but enabled by the technology. This new science environment demanded separate recognition. Wulf named it the “National Collaboratory.”

The National Collaboratory -- White Paper (Wulf 1988)

Since 1988, William A. Wulf has been the AT&T Professor of Engineering and Applied Science at the University of Virginia. Wulf concentrates on undergraduate computer science education, research on computer architecture and computer security, and assisting humanities scholars in the exploitation of information technology. From 1988 to 1990, Wulf served as an Assistant Director of the National Science Foundation, specifically as the Director of the NSF Directorate for Computer and Information Science and Engineering (CISE).

In May 1998 he wrote his widely-cited but still largely unavailable *White Paper*. It was not published nor made publicly available.

In March 1989, Wulf convened a select group of twenty-nine researchers and scientists for an invitational workshop at Rockefeller University, and presumably read or otherwise delivered the *White Paper's* message. The convention produced Lederberg and Uncapher's (1989) report, *Towards a National Collaboratory*, which includes the *White Paper* as Appendix A. The Lederberg and Uncapher report was also not published. Neither document is catalogued as held by any lending library in the world, nor are they available from the National Science Foundation, or from the authors. Because the *White Paper's* "center without walls" quote is perhaps the most frequently repeated and often miscited passage within the collaboratory literature, and because neither of the documents was ever published, an in-depth look at both of them is warranted.

The Wulf's *White Paper* is a two-page, single-spaced document with two sections: "Background" and "The Proposal." In the Background section, Wulf contends that

The health of the United States, economically and militarily, depends on technology... [which] ...depends on the number, quality, and productivity of the nation's research scientists and engineers... [which] ...depends on such things as adequate facilities, stimulating colleagues, and the open exchange of ideas. (1)

Wulf describes the emergence of geographically situated interdisciplinary centers, institutes, and laboratories which had already produced a "disproportionate share of the advances of their respective fields" (1). He describes the coming trend in which such centers will no longer be geographically determined, but will be "freed from the constraints of distance" producing research teams for which "opportunity and choice will determine the composition, size, and duration" (1).

Wulf warns that it will no longer be necessary for such centers to share a common administrative structure. He foretold a fundamental shift in the way science is conducted. Remote interaction with instruments, colleagues, and data will not only be possible, but mandatory. Interaction with remote instruments will be necessary either because the instruments are too expensive to be widely held (space telescopes), or the environment in which they function is inhospitable to humans (deep ocean vehicles). Interaction with remote colleagues will be mandatory because the talents necessary to address interdisciplinary problems will not be collected in any one place. Remote access to data will be necessary because the data is too vast to be replicated. Finally, some of the most pressing scientific challenges facing us, such as that of the global change, are inherently distributed and exhibit all of these properties; remote interaction with instruments, colleagues, and data is essential to solving them. (1)

Wulf proposed a

major, coordinated program of research and development leading to an electronic 'collaboratory', a 'center without walls', in which the nation's researchers can perform their research without regard to geographic location—interacting with colleagues, accessing instrumentation, sharing data and computational resources, accessing information in digital libraries. (1)

The enabling technologies, Wulf wrote, are "high speed information processing and communication" (1). By 1988, a national research network was already underway; high performance computers were becoming ubiquitous; and open system standards were making these facilities more accessible to the research community.

In a sense, the collaboratory was an inevitable outcome of these developments. However, much as a 'center' is enabled by the building that houses it but is not just the building, the collaboratory is not just interconnected computers.

A complete infrastructure is required: software that facilitates collaboration, simulation tools that can substitute for some aspects of the traditional wet laboratory, 'smart instruments' that can be used effectively remotely and interchangeably [sic] with simulated experiments, digital libraries and software to access the information in them, accessible (usable) repositories of raw data, etc. (1)

A great deal of research still needed to be done "to exploit the enabling technologies and build this infrastructure" (2). Some of the research, Wulf said, was needed in the areas of traditional computer science and

computer/communication engineering (network speed, security and integrity of communication, smart instrumentation), but some of the research was essentially social, behavioral, or economic:

How do people collaborate, and how can we exploit technology to amplify the effectiveness of this collaboration, especially when the collaborators are not colocated? (2)

Wulf explained that while much of this research was underway, it was not coordinated and hence was not easily combined; that

when viewed independently, aspects crucial to the total concept are not perceived to have an especially high priority. By setting ourselves a concrete goal we can focus the energies of the research community in a way that will ameliorate both of these problems. (2)

Wulf asked workshop participants to imagine the impact the collaboratory would have on the productivity of the nation's scarce human resources. He argued that while it could be said that the collaboratory "would not enable anything new—anything not now possible, albeit more slowly or at the cost of moving people..." that, "the quantitative increase in ease of collaboration will have a profound qualitative effect" (2). That qualitative effect would:

- enhance the productivity of the individual researcher by providing access to information and instrumentation now available only at prohibitive cost in both time and money.
- increase the number of nimble minds and diverse perspectives far beyond those available at the researcher's home institution.

- enable both inter- and intra-disciplinary research that simply isn't being done now because the best people to do the research are not collocated and the scale or duration of the project does not justify a megacenter and the associated relocation of people.
- increase the pool of researchers available to work on the problem. The faculty at four-year and predominately minority institutions are an essentially untapped resource. The collaboratory will permit them to be full and effective partners in research projects, and increase the quality of instruction at those schools at the same time.
- speed the transition of new ideas into industry, into products—and increase the relevance of research to social/economic goals. By making academics and advanced developers in industry a part of the same collaboratory, the same 'intellectual stew,' we can achieve both effects simultaneously and naturally.

Wulf warned that we might never know quantitatively the impact of these combined effects because

we don't know how to define 'research productivity' quantitatively, but also because we may not know what would have happened without the collaboratory. Moreover, not every researcher will wish to collaborate remotely, nor will any one single technological 'fix' cure the myriad problems faced by the country. (2)

Nevertheless, Wulf proposed, the effects will be profound. He explained that just one of the impacts outlined, the speed of technology transfer from research idea to product, had already been shortened from the usual 15-20 years to less than four years.

So, Wulf asked, "What needs to be done?"

For the most part we do not need to begin whole new areas of research. Rather, we need to coordinate and expand research already underway, to deploy the enabling infrastructure...to guide the establishment of national projects (such as the human genome and global-change databases) along paths that will permit them to interoperate, and, eventually, to set standards for both commercially-produced and one-of-a-kind instrumentation to be usable remotely. (2)

The first step, coordinating and expanding research underway, was to be undertaken at the Rockefeller University workshop. Wulf charged the workshop with setting the research agenda for the collaboratory, by addressing such questions as

- What are the central problems?
- What is the best mode for attacking each of them?
- How can research be coordinated?
- Who are good candidates for demonstration projects?
- How best can results be used as they emerge? (2)

Wulf concluded his *White Paper* with the recommendations that the NSF play a leadership roll implementing the agenda, but warned that the effort may become,

in both size and scope, larger than any one agency can handle and will also require heavy involvement of the private sector (2).

Towards a National Collaboratory (Lederberg and Uncapher 1989)

The goal is to build no less than a distributed intelligence, fully and seamlessly networked, with fully supported computational assistance designed to increase the pace and quality of discourse, and a broadening of the awareness of discover: in a word, a collaboratory. (Lederberg and Uncapher 1989, 3)

The report of the Invitational Workshop Wulf convened at Rockefeller University on March 13-15, 1989, *Towards a National Collaboratory*, endorsed the concept of the collaboratory enthusiastically. It recommends a "three-fold agenda that would repeat itself in cycles of design, implementation, and testing" (8). The first cycle concerned systems architecture and integration, and would examine ways to allow people and machines to use the collaboratory's components most effectively. The second cycle would evolve tools and technologies themselves, and the third would develop user-oriented testbeds to validate both technology and organization.

Lederberg and Uncapher's 18-page report has five sections with an Introduction. The sections are:

Science and Support for Collaboration in which the concept of team science is discussed within the framework of Lederberg's "Epicycles of Scientific Discovery" model (Lederberg 1989).

Expected Impact in which the collaboratory's potential amplification of future research is discussed with particular attention paid to the impacts on:

- Global Change research underway in the NSF, NASA, United States Geological Survey (USGS), and Department of Energy (DOE);
- The Human Genome project underway at NSF, The National Institutes of Health (NIH), and DOE; and
- Parallel Processing Research underway at NSF, DARPA, NASA, and DOE.

The Functional Collaboratory – which provides a functional description of, and assumptions about, the collaboratory.

Research Agenda in which details of the three-fold approach are provided.

Conclusions and Recommendations, which details the critical factors and issues to which attention must be paid.

This study examines the Functional Collaboratory, Research Agenda, and Conclusions and Recommendations sections of the Lederberg and Uncapher report in detail.

The Functional Collaboratory

What is a collaboratory? As we use the term here, it is the combination of technology, tools and infrastructure that allows scientists to work with remote facilities (co-laboratory) and each other (collaborat-ory) as if they were co-located and effectively interfaced. (Lederberg and Uncapher 1989, 6)

First, the authors warn, it is important to remember that scientists have always collaborated. Such collaborations include the processes of publication, co-authoring articles and attending conferences, sharing students, and working on teams at experimental facilities.

The norm of mutual scientific criticism is an intense form of intellectual collaboration, however antagonistic it may appear. Besides its cognitive utility, criticism is also indispensable to a rational system for the allocation of resources, tenured positions, research funds, and facilities. (6)

As science becomes more multidisciplinary and requires increased access to expensive and rare resources that are impossible to duplicate, the collaboratory will allow access to this remote instrumentation, "greatly expediting such scientific research: The collaboratory will provide seamless access to colleagues, instruments, data, information, and knowledge" (6).

Lederberg and Uncapher's (1989) report provides the intellectual foundation of the collaboratory. It relies on two distinct models: The first model is Lederberg's "Epicycles of Scientific Discovery," which outlines the processes and phases that can and should be supported through the collaboratory. The second

model is Mark Stefik's "The Functional Collaboratory", a flowchart that shows how each of Lederberg's research functions might be supported through the technical capabilities in the collaboratory.

For example, one critical function of scientific research, project organization and management, is examined. The functions required from a collaboratory to support coordination of action, joint design, and resource scheduling would include

email and directories, tools to support structured discussions, a digital library with appropriate search mechanism, user education and training tools, real-time computer supported multi-media teleconferencing, a remote experiment scheduler, and so on. These capabilities are in turn supported by a number of enabling technologies, such as networks, advanced human-machine interfaces, high resolution displays, and video compression techniques. Finally, underlying the collaboratory paradigm must be an understanding of how groups work together as well as offering seamless new technologies to enhance the availability of old knowledge, permit scientists new means of accelerating the pace of discovery and support the amplification of human intellectual capability. (7)

The Research Agenda

The recommendations for a research agenda outlined in Section Four of the Lederberg and Uncapher report are based on the following assumptions:

- The underlying communications network provides a high level of minimum capability (TI communication lines with 1.544 Mbps throughput and

sufficient bandwidth [T3 or 45Mbps] to support anticipated load without delay; eventually achieving 1 Gbps throughput);

- Computing systems with relatively high capability (workstations with 10 Mips processor speed, 10 Mbytes memory, and 1000x1000 pixel color display connected to high performance computing with processor power of gigaflops and beyond;
- Infrastructure described above is made available to users/scientists. (7)

The research agenda focuses "on developing and demonstrating the technologies required to make the National Collaboratory a reality (8). The agenda recommends a three-fold approach:

1. Systems architecture and integration aimed at the system issues that allow people and machines to effectively use various technologies that are the components of the collaboratory;
2. Evolution of the underlying technologies and tools themselves; and
3. User-oriented testbeds coupled with better theories about the process of collaboration that are required to validate both the technical approaches and the overall system components allowing understanding of the requirements on such a system and the role that the various technological components play in the overall design.

The three-fold approach is envisioned as "an iterative cycle of design, implementation, and testing" (8).

Systems Architecture

The report admits that most of the research that needs to be done to actualize the collaboratory is in the area of system architecture and integration, but it echoes Wulf that the collaboratory is more than a set of tools.

It is a functional capability to improve scientific effectiveness by taking advantage of a broad set of resources, including but not limited to remote facilities and other scientists. Understanding the appropriate system architecture, where such architecture includes not only the underlying tools but also the people that are to use those tools, requires a dedicated effort. (8)

Available Technologies

Lederberg and Uncapher point out that there were already tools available, both from the research community and the commercial sector, on which to begin design research on the systems architecture outlined above. These tools (and their limits, and suggested areas for research) include:

Electronic Mail (which needs interoperability, graphics capability, privacy, and user support);

Electronic File Transfer, which was already well proven;

Remote Access and Control, including remote logon (in need of enhancements to assure access control and authentication for safety and security);

Shared Files (which allows the sharing of ASCII files themselves, but with limited ability to share information through such files, and in need of a higher level of functionality and standardization to include graphics and data research representations);

Database Access (ability to store and retrieve data from shared databases needs standardization);

Access Control and Authentication (security mechanism need to be adopted and refined);

Multimedia Mail (integration of graphics, sounds, spread sheets, scanned images into text requires powerful mechanism for interoperability);

Structured Interaction Support (migration of multi-media teleconferencing, structured conversations, and information sharing from proprietary platforms to open architecture);

Simulation of Instruments (prototyping instruments through computer simulations via totally compatible hardware and software with remote debugging capabilities; after which, the distribution of simulated

instruments to large number of scientists for conduct of large, simulated distributed experiments).

The recommended research agenda would investigate the integration of these tools and the development of user interfaces in a context of user feedback (with the following objectives):

Integrating Technologies

Digital Instrumentation (dedicated to theories and technologies for development of real-time control and feedback from instruments, and real-time communication concerned with remote and multi-user instrumentation, experiments with communication and control technologies, development of interfaces to address delays and scheduling, and use of human and machine agents);

Multi-media Meetings (experiments on collaboration in meetings, distributed education; economic means for providing very high-bandwidth transmissions; social experiments to study the effects of technology on interpersonal argumentation at a distance);

Digital Mail (develop ability to send a verified and trustable electronic check, linking value-added services, better addressing mechanisms, comprehensive yellow-page services, technical extensions and social issues surrounding the use of electronic mail);

Scientific Reference Service (service to provide expertise and network to answer tough questions, a “Who Knows...?” service using literature access and intelligent agents, but drawing first on the human, then on artificial intelligence);

Digital Journals and Peer Review (services for logging documents, logging comments on documents, and offering digital document retrieval; support for collaborative writing, experimentation with different modes for commentary, editing, and document exchange; social experiments to determine salient effects on the perceived qualities of number of reviews, and effects on peer group size);

Digital Library (many variations including software, video, and other “unusual forms” in a distributed electronic database; collaboration and electronic publishing; integration of services over distributed libraries; social experiments with policies, citation mechanisms, pricing, collection and distribution of royalties);

The digital library was given added attention in the report’s section on integration of technologies. Of particular interest was development of techniques for discovery through digital search, including technologies for scanning and comparing strings of digital data; document storage; and content recognition capabilities through large-scale linguistic analysis or comparison. Social

questions about the digital library include analysis of the ways that automated searching enable collaboration or discovery. "The digital library is likely to be the key to valuable old knowledge, and new knowledge so vital to the scientific process" (12).

Advanced Technologies

Developing the integrated capabilities outlined above, the report says, will require developing tools that did not yet exist. These underlying technologies are described as:

Hypermedia Conversion Support such as hypermedia databases to track design decisions, operational problems and corrections, and research approaches.

Intelligent Agents such as distributed processes that would act on behalf of users. The report visualized each entity in the collaboratory (scientists, instruments, databases, computer resources) having an intelligent agent between it and the network. These intelligent agents would act on behalf of the entity, negotiating with other agents, conducting searches, scheduling, etc.

Interoperable Data Description that would describe data from multiple disciplines using a common format to allow interoperable data analysis and manipulation.

Information Fusion that would allow integration of information from heterogeneous sources.

Smart Agents for the Design of Experiments to facilitate the use of multi-sensor experiments by multiple investigators.

Smart Data Gathering incorporating intelligence into the instrument that would allow “self-directed” data gathering (13).

Technology Utilization

How the emergent technology is used, and “the relationship between technologies and the way scientists do and will conduct their research” (14) is identified as a critical issue needing investigation in order to make the collaboratory a reality. Lederberg and Uncapher recommend that appropriate testbeds to understand the impact of technology be coupled with a research program into the underlying mechanisms of the collaboratory itself.

User Testbeds

Historically, many new technologies have been left dangling at the end of the research cycle waiting to be adopted by some user community or integrated into commercial projects (14). Lederberg and Uncapher suggest setting up user-oriented rapid-prototyping testbeds as partnerships between users and developers, and identified several critical attributes such testbeds must have:

- Represent a partnership between users who see the potential for the collaboratory improving their scientific research and technologists who are interested in working with such a community.
- Be of a size that is sufficient to explore the impact of the collaboratory on scientific research (which means a team of geographically dispersed users working with laboratory facilities) but small enough to be manageable as a rapid prototyping environment.
- Provisioned with adequate infrastructure (networking, workstations) so that the prototype does in fact represent future potential.

The report notes that it is important that those who are working on developing technologies be provisioned at any stage with the next stage of infrastructure so as to act as a leading edge for the technology.

Collaboration Mechanism

In order to develop tools for supporting scientific collaboration, the Lederberg and Uncapher report insists it is critical to understand the process of collaboration itself. Achieving this understanding is a multidisciplinary enterprise and will draw on a variety of existing disciplines. Some of the work will be observational (field studies, surveys, archival analysis), some will be experimental. The testbeds will be one environment for conducting this research.

It is also crucial to synthesize previous work into new theories about how collaboration and coordination occur and how technology can help (15).

Examples of Research Questions

The Lederberg and Uncapher report provides a description of some of the questions that needed to be addressed by the National Collaboratory research agenda. Those questions include:

What are the basic processes involved in coordination?

What structures are possible for carrying out these processes?

How is collaboration among scientists different from other kinds of collaboration?

How might use of collaboration technologies affect incentive structures for the conduct of science?

Is the social science structure of science affected by intensive use of communication technology for remote collaboration and resource sharing?

To what degree is it possible to substitute capital (electronics) for scientific labor?

The Report's Conclusions and Recommendations

A number of factors and issues addressed in the conclusion of the Lederberg and Uncapher report bring firm focus to the critical elements necessary for developing the National Collaboratory:

- Importance of integration and user testbedding
- Careful selection of user testbed communities
- Community workshop to bring users and developers together
- Targeted integrated system that all works together
- Drawing on experience of other agencies/organizations

Conclusion

The foundation of the collaboratory is provided by three key documents. The first two of those documents were never published nor widely distributed, although they are widely cited, and are examined in detail in this chapter. The first document, William Wulf's *White Paper* (1988) puts forth the concept of the collaboratory, sets its philosophical foundation, and identifies the disciplinary and research focus needs of the National Collaboratory. Wulf projected that the collaboratory would be an interdisciplinary endeavor requiring a relatively equal contribution from a variety of disciplines.

The second key document, Lederberg and Uncapher's (1989) *Towards a National Collaboratory*, sets the intellectual foundation of the collaboratory, outlines the National Science Foundation's National Collaboratory research agenda, and identifies the topics and approaches of research needed. Lederberg and Uncapher reflect Wulf's suggestion that the collaboratory be constructed as an interdisciplinary environment from relatively equal contribution from multiple disciplines.

Chapters Three and Four of this study constructs taxonomies, or categorical classification schemes, based on the Wulf and the Lederberg and Uncapher documents. Chapter Three also constructs a taxonomy based on Haddow's (1997) types of publication. These three taxonomies are used to guide exploration of the collaboratory literature, to conduct taxono-bibliometric analysis of the literature to determine the extent to which the assumptions of relative equality of contribution to and interdisciplinarity of the collaboratory environment are reflected in the subsequently published works of collaboratory researchers, and to arrive at an objective reality of the collaboratory. The third key document, the National Research Council's *National Collaboratories: Applying Information Technology for Scientific Research* (1993) is examined in Phase Two of this study.

CHAPTER THREE

Discipline, Focus, and Type of Publications

Wulf invented the word "collaboratory" "to combine the words collaboration and laboratory," (2) and to define a

'center without walls,' in which the nation's researchers can perform their research without regard to geographic location—interacting with colleagues, accessing instrumentation, sharing data and computational resources, accessing information in digital libraries.(1)

Wulf's (1988) *White Paper* identifies the disciplines that needed to participate in collaboratory research, and the focus that research needed to take. Chapter Three constructs a taxonomy based on Wulf's disciplines and focus, and tests if the library literature reflects the relative equality of contribution to, and interdisciplinarity of, the collaboratory environment he assumed.

The Document Retrieval Set

Eighty-seven documents containing the word "collaboratory" were retrieved during an exhaustive search of the library resources available to students and faculty through the Mary Evelyn Blagg-Huey Library at Texas Woman's University. Texas Woman's University is a Carnegie I Comprehensive Land

Grant State University. In addition to extensive book and serial holdings, the Blagg-Huey Library provides access to over a hundred proprietary electronic databases, and on-site CD-ROMs. Many of the articles retrieved were represented in several resources.

The search was difficult because there are no established journals, print or electronic, nor are there documentary portals to the collaboratory literature; therefore every resource had to be searched individually in order to achieve a comprehensive retrieval set. The search was further complicated because the word "collaboratory" is not included in any database thesaurus, nor is it a subject heading in indices. Collaboratory is not a Library of Congress Subject Heading.

In the case of bibliographic searches, the cataloger or abstracter would have had to use the word specifically, or it would have had to appear in the title of the work in order for a record to be retrieved. In the case of fulltext search, the word could also have appeared in the body of the document. Because "collaboratory" was coined by Wulf in 1988, all the documents contained in the retrieval set were published, cataloged, and entered into the library holdings between January 1, 1988 and December 31, 1998, and thus represent the documentary evidence of the first ten years of the collaboratory. The two key documents explored in Chapter Two never appeared in the retrieval sets.

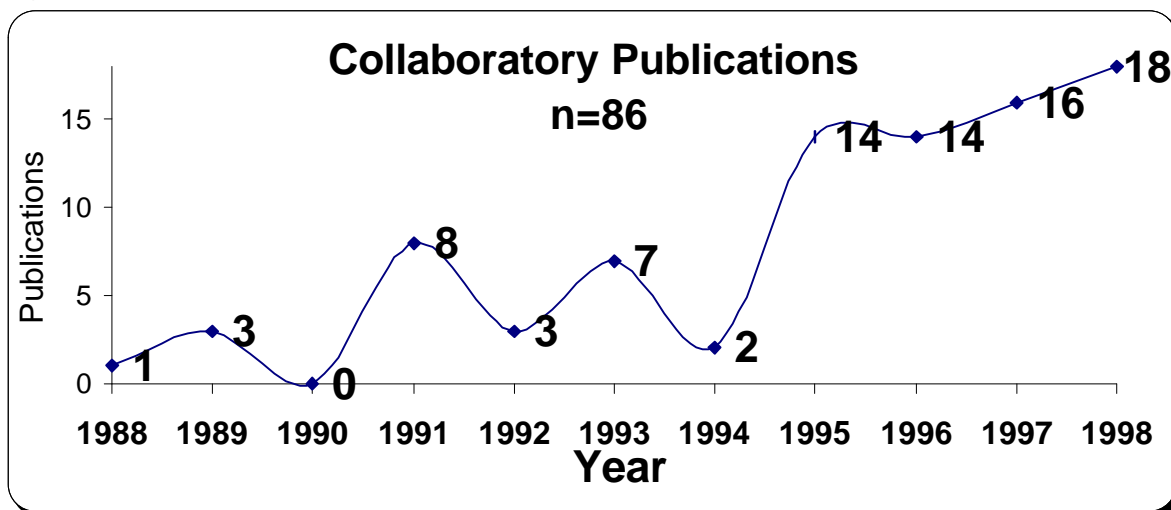
The third key document, NRC's (1993) *National Collaboratories* (which is discussed in Phase Two of this study) was removed from the retrieval set for use along with Wulf, and Lederberg and Uncapher, during analysis of the remaining eighty-six documents. The eighty-six documents were indexed and abstracted into an Microsoft Access97® database. Admittedly, the document set by count forbids analysis using statistical inference. However, the retrieval set is important as it is comprehensive, and thus highly representative of the documentary evidence of the first decade of the collaboratory. The library documents are the only publicly preserved record of the collaboratory, and thus their compilation here has historical value in addition to their value to this study. The eighty-six documents are cited and annotated in Appendix A.

The proprietary databases searched included, but were not limited to, FirstSearch™, Dialog™, Ebsco™, ERIC™, GaleNet™, Academic Abstracts™, Articles1st™, General Science Abstracts™, Dissertation Abstracts™, Social Science Abstracts™, OCLC™, and Library of Congress Catalog. The search and retrieval activity took place during weekly in-library and frequent extra-library sessions between August 1998 and December 1998, and consumed well over 100 person hours. The search and retrieval frequently relied on the advice and assistance of professional librarians, particularly for retrieval of many of the documents via Interlibrary Loan.

The retrieval set represents a document base that any serious and library-savvy researcher conducting traditional “library research” and exploiting the full capability of the library, could achieve. The retrieval set does not assume that every relevant or pertinent document was located and retrieved, but it does presume to be highly representative of the available literature and therefore a proper foundation on which to build an objective reality of the collaboratory.

Figure 1 shows the publication dates of the eighty-six documents.

Figure 1. Publications by Year



Taxonomy Construction

A taxonomy is an analysis tool that facilitates identification, naming, and classification of objects. For this study, three taxonomies are constructed. Each taxonomy has two main categories. Each main category has subcategories. The taxonomies are constructed using Microsoft Excel97® spreadsheet software, with the main categories placed on the X- and Y-axes. The subcategories are placed in columns along the X-axis and rows along the Y-axis, as shown in Table 1.

Table 1. Basic Taxonomy

		X Axis		
		X Subcategory	X Subcategory	X Subcategory
Y Axis	Y Subcategory	Cell 1	Cell 4	Cell 7
	Y Subcategory	Cell 2	Cell 5	Cell 8
	Y Subcategory	Cell 3	Cell 6	Cell 9

The intersection of each column and row is a cell in the taxonomy. Each cell is assigned a distinct number. Documents are evaluated and assigned to cells according to their best fit in the X and Y categories and subcategories. Cell

numbers are then used to analyze the document set. The categories and subcategories for each taxonomy are discussed in detail in the following sections.

The eighty-six documents in the retrieval set were examined individually as they were retrieved and entered into the database, then analyzed as a body three separate times, once for each taxonomy. Each document is categorized by assigning the number from the most representative cell in each taxonomy. The cell numbers were entered into corresponding fields created in the Microsoft Access97® document database. The database fields were imported into Microsoft Excel97® for statistical analysis.

The Wulf Taxonomy and Analysis

Wulf's (1989) *White Paper* identifies three discipline classes that needed to contribute to research about the collaboratory:

- Computer Science (CS)
- Computer/Communication Engineering (CCE)
- Social, Behavioral, Economics (SBE)

and three areas that needed research focus:

- Instrumentation
- Colleagues

- Data

A 3x3 taxonomy based on Wulf's disciplines and focus was drafted, and an analysis of the document retrieval set undertaken. For multiple-author publications, the discipline of the lead author was used. The discipline of the lead author was determined by information provided with the article, usually an affiliation in the byline, author biographical note, or email address. In the event the primary affiliation of the author was not determinable, the article was placed in the most logical category based on the journal of publication or topic of research.

It became quickly apparent it would be impossible to maintain precise distinction between the first two of Wulf's discipline categories, so the Computer Science and Computer/Communications Engineering categories were collapsed into a single category, CS/CCE. Two additional categories emerged during document analysis and coding, and were created. The first new category, Library Information Science (LIS) was added because of researcher preference and curiosity. The second new category, Other, was added because many of the disciplines did not fit the Wulf categories neatly. Library Information Science includes authors whose primary affiliation is either a library or information agency, library school, or an information school or department. Disciplines assigned to the Other category includes authors from government, journalism,

medicine, education, chemistry, botany, physics, biology, and chemistry, among others. Abbreviations for the revised discipline categories are provided in Table 2. The eighty-six documents were categorized into the reconstructed 4x3 Wulf taxonomy without problem, and a second coding was performed to assure consistency.

Table 2. Taxonomy of Disciplines

CS/CCE	Computer Science, Computer Communication Engineering
SBE	Social, Behavioral, Economics
LIS	Library Information Science
Other	Government, Journalism, Physics, Medicine, and others

Several articles had multiple foci. The primary focus of each article was used. Articles concerned with hardware, software development, or infrastructure were placed in the “Instrumentation” category. Articles with a primary focus on humans or human groups were placed in the “Colleagues” category. Articles with a primary focus on data (acquisition, manipulation, data sharing, standards, etc.) or information (storage and retrieval, search algorithms, metalanguages) were placed in the “Data” category.

To clarify how specific articles might be categorized, a library-accessible document that reports research performed by a computer communications

engineer, and which is about the development of a routing protocol that allows multiple simultaneous occupants of a singular documentary space over computer networks, and retrieved using the search string "collaboratory" would be placed in the CS/CCE x Instrumentation cell. The cell was assigned an address of one, so that document, and all others fitting that category, would be assigned a taxonomic code of one. A paper published by a sociologist addressing the impact of that interface on successful conduct of international business communications within the collaboratory would be placed in the SBE x Colleagues cell, which is assigned the taxonomic code six.

Each document's taxonomic code was entered in the document database, then imported to a spreadsheet. Counts and subtotals were calculated. Table 3 shows the number of documents assigned to each of the cells within the discipline by focus taxonomy.

Table 3. Publications by Discipline and Focus

Disciplines	Focus of Research			Totals
	Instrumentation	Colleagues	Data	
CS/CCE	30	2	2	34
SBE	14	4	0	18
LIS	5	2	1	8
OTHER	18	6	2	28
Totals	67	14	5	86

Computer Science/Computer Communication Engineering produced thirty of the sixty-seven Instrumentation publications, the largest category. CS/CEE also produced two of the fourteen Colleagues publications, and two of the five publications concerned with Data, for a total of thirty-four publications.

Social, Behavioral, or Economics (SBE) produced eighteen publications; Library Information Science (LIS) produced eight, and Other disciplines produced twenty-eight. Figure 2 shows the percentage of publications, by discipline, and Figure 3 shows distribution within the full discipline by focus taxonomy.

Figure 2. Percentage of Publications by Discipline

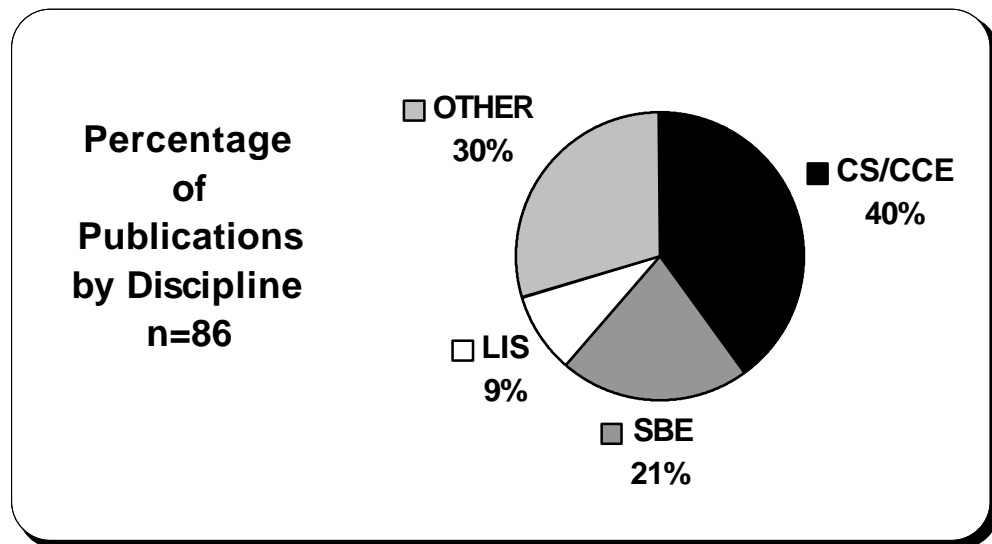
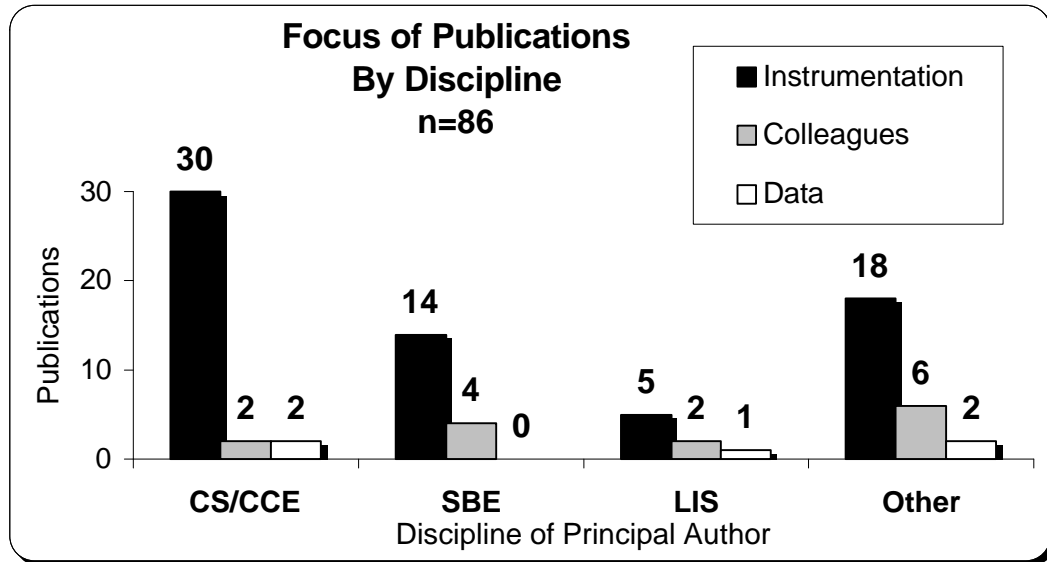


Figure 3. Publications by Discipline and Focus



The "Other" category of disciplines produced 30%, or twenty-six of the eighty-six documents, including eighteen Instrumentation documents, six Colleague documents, and two Data documents. The number and percentage of Other disciplines confirms Wulf's assumption that the collaboratory would be an inherently interdisciplinary environment constructed by multiple disciplines. Because the Other category was added during data analysis and contains a significant number of articles, a second taxonomy based on Haddow's (1997) types of publications is constructed.

The Haddow Taxonomy and Analysis

Many of the documents in the eighty-six document retrieval set were not research articles in the traditional sense. Haddow's (1997) review of the nature of journals of librarianship offers a second taxonomy into which the collaboratory articles are placed for a more exact contextual analysis.

Haddow relied on Price, and Windsor and Windsor's definition of 'scholarly' literature as hinging on the number of citations an article includes, with those publications having citations classified as scholarly, and those without citations not scholarly. She recalls Beal's description of library publications as either "glad tidings, testimony, or research," type articles, and Saracevic and Perk's further classification of "news-type articles."

For this analysis, several of Haddow's (1997) classifications were combined. Articles concerned with visions, speculations, and success stories, and which contained no or minimal citations, are placed in the "Glad Tidings and Testimony" category. Articles with substantial citations, published in scholarly journals, that develop theory or hypothesis or provide detailed analysis, development, or implementation details are placed in the "Research" category. Articles concerning announcements of funding, partnership formations, commercial implementations or software or hardware rollouts and the like, are placed in the

"News-type" category. Table 4 illustrates the placement of the articles in the retrieval set within the article type taxonomy.

Table 4. Publications by Type

Type of Article	Number of Articles
Glad Tidings and Testimony	14
Research	22
News-Type	50
Total	86

Fifty of the articles, or 58% of the total, were placed in the News-Type category. Fourteen articles, or 16% of the total, were placed in the Glad Tidings and Testimonies category. Together, the News-Type and Glad Tidings and Testimony categories represent 74% of the retrieval set. Figure 4 shows the number publications by type of article. Twenty-two of the eighty-six publications, or 26% of the collaboratory literature are classified as Research. The publication types are plotted by publication year in Figure 5.

Figure 4. Article Type

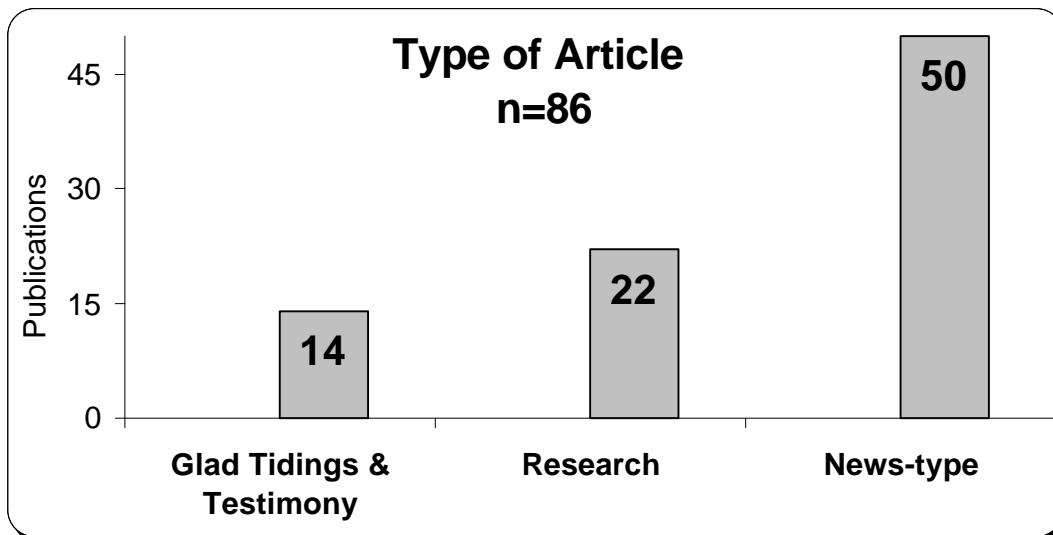
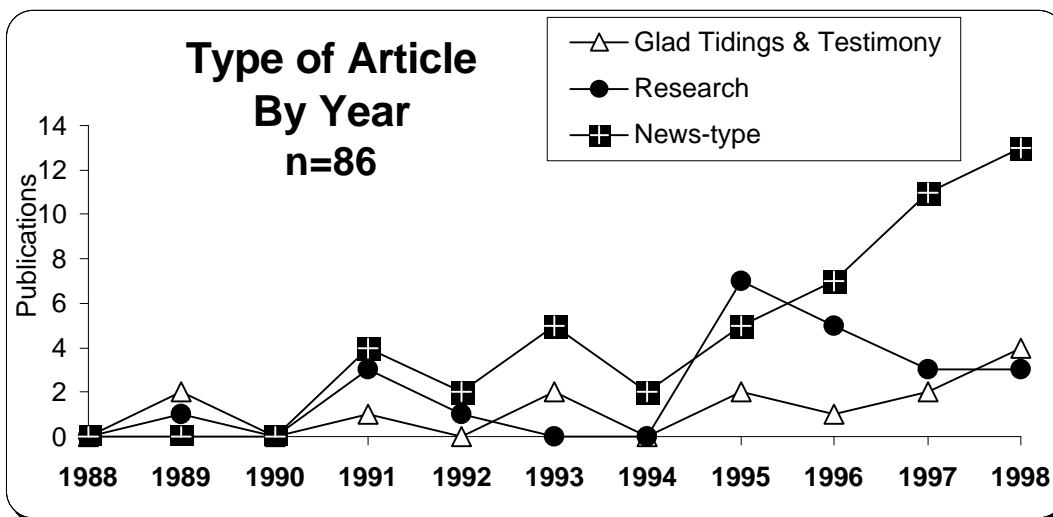


Figure 5. Article Type and Year



Two Glad Tidings articles and the collaboratory's first Research article were published in 1989, the year Wulf's *White Paper* was presented at the Rockefeller Workshop. There were no publications in 1990. The first official collaboratory-specific federal research funding was announced in 1990, and a surge of News-Type articles followed. News-Type publications dominate the collaboratory literature for the next four years, until 1995, when, for the first and only time in the ten year history of the collaboratory, more Research articles were published than any other type.

News-Type publications began four consecutive years of increased frequency in 1994, reaching a high of thirteen News-Type articles in 1998, a year that also saw the most number of Glad Tidings and Testimony articles. Quantitatively speaking, the seven research articles published in 1995 mark the research high-point of the decade. The year 1995 also marks the beginning of an overall upswing in the total number of publications, with a peak 20 articles published in 1998, the last year included in this study.

Figure 5 draws attention to three two-year "waves" in publications: 1988-1990, 1990-1992, 1992-1994. These waves roughly correlate with announcements of funding for collaboratory research. These early indicators of trends, waves, and possible correlation in the types of articles indicates a potential research agenda interested in analyzing media effects, or publication

affects in relation to federal funding, or by other aspects of mass media, mass communication, or political science theories.

Wulf noted in his *White Paper* that the research leading to the collaboratory is uncoordinated. So is the federal government's funding documents related to collaboratory research. Full fiscal documents could not be retrieved during the five-month data gathering period of this study because, like the research documents, they are buried here and there in the volumes of data, and use an uncoordinated vocabulary with undeveloped access portals. For example, of the forty-one NSF projects funded in the 1998 Knowledge and Distributed Intelligence Initiative (KDI), a \$350 million, seven-year research program intended to specifically advance research about the collaboratory, only one funded proposal uses the word "collaboratory" in its title.² A fulltext fielded search of the abstracts of the NSF funded research database for keyword "collaboratory" produces seventeen funded projects, dating from 1990, including the one described above.³

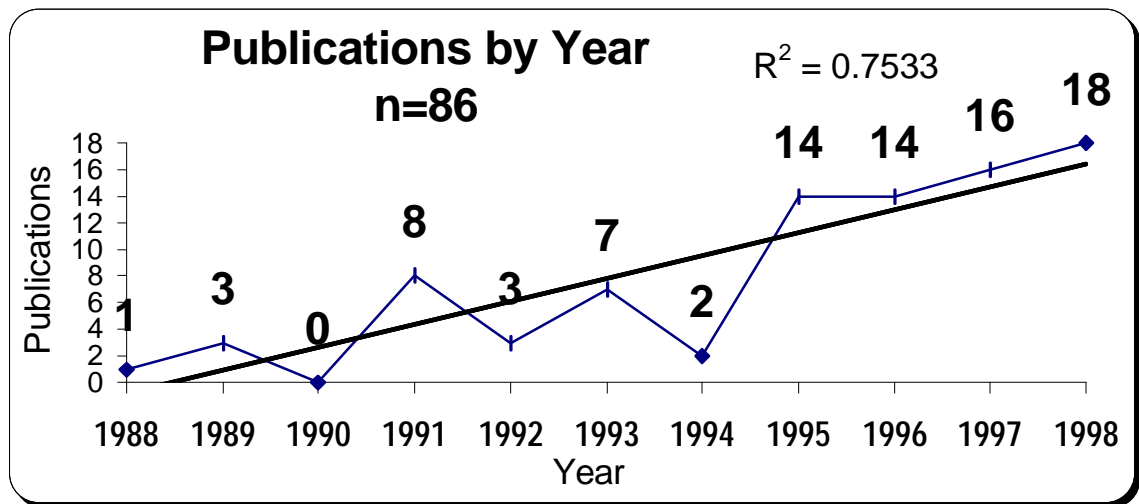
² See an html-ized version of the NSF Comma-separated-value (CSV) concatenated dataset of 1999 KDI awards at www.intertwining.org/collaboratory/1998_Awards.htm .

³ The search page for NSF-funded research projects is at <http://www.nsf.gov/verity/srchawdf.htm> . The KDI webpage is at <http://www.ehr.nsf.gov/kdi/>

Descriptive Statistical Analysis

Figure 6 shows a moderately strong positive linear trend ($R^2 = 0.7533$) in the number of collaboratory publications over time. Granted, the dataset is too small for rigorous inference. Nevertheless, it is clear that the concept of the collaboratory, as reflected by a decade of representative literature, has sustained positive interest.

Figure 6. Trends Analysis, Publications over Time



Many of the fifty News-Type publications identified by the Articles Type taxonomy were written by or about individual collaboratory participants who are actively engaged in scholarly or scientific research, and most are clearly attempts

to foster popular understanding and garner popular support for the notion of the collaboratory. Simultaneous with the rise in News-Type publications in the last half of the decade, there was a rise in Glad Tidings and Testimony publications but a decrease in the number of Research articles. An average of one Glad Tidings and Testimony article was published each year. Many of these Glad Tidings and Testimony articles were also written by or about a core group of researchers or projects, indicating a concerted public relations effort. Future analysis might include publication production by author or project, and citation and co-citation analysis to reveal the "personalities" of the collaboratory.

Discipline x Focus x Article Type Co-analysis

This section analyzes the twenty-two research-type articles identified by the Haddow Taxonomy by focus and discipline of the Wulf Taxonomy. Figure 7 shows the subset of research-type articles by research focus. Fifteen of the twenty-two research-type articles focus on Implementation, six focus on Colleagues, and one is about Data.

Figure 7. Focus of Research Articles

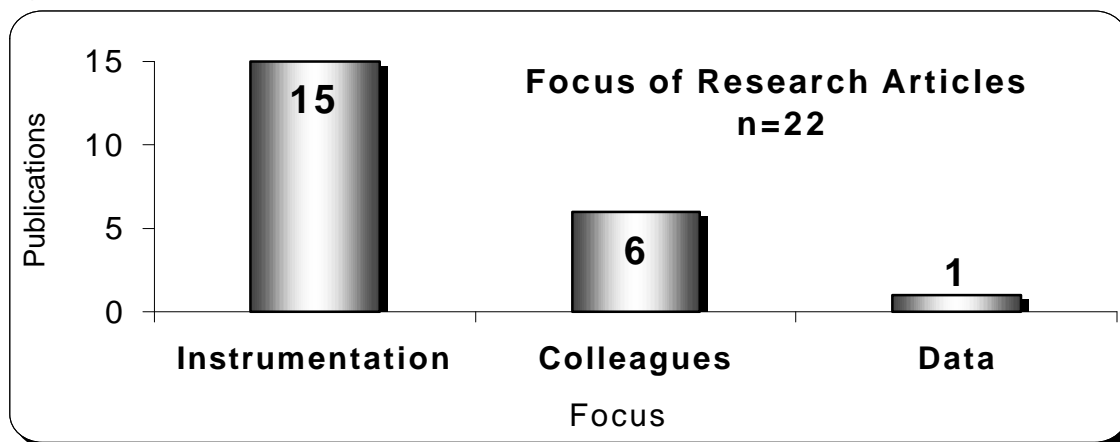
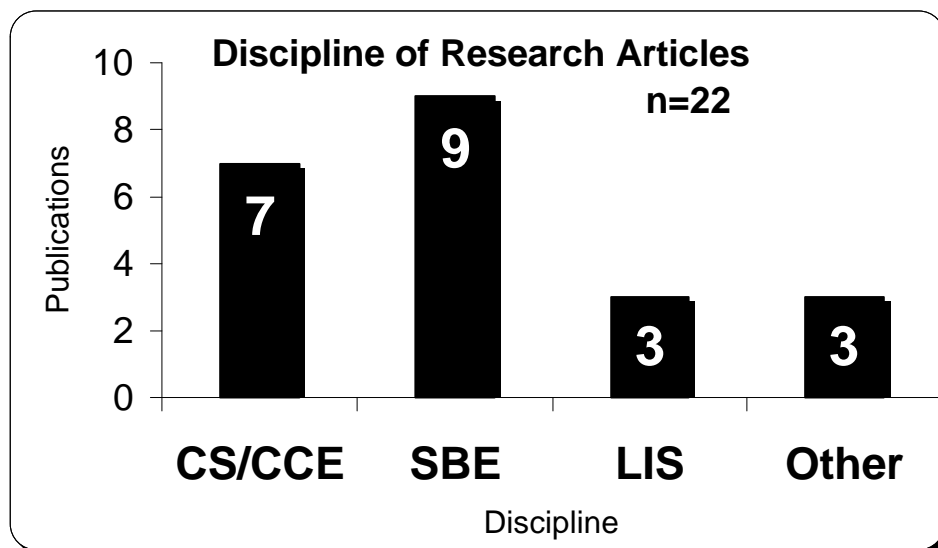


Figure 8 shows the twenty-two research-type articles by discipline of the primary author. Seven of the articles, or 32%, originate from Computer Science/Computer Communication Engineering (CS/CCE). Nine of the research-type articles, the largest portion, are published by the Social, Behavioral, and Economics disciplines (SBE). Library Information Science (LIS) produced three of the twenty-two research-type articles, and Other disciplines produced three research-type articles. It is likely that at least some of the articles were co-authored by scholars from different disciplines.

Figure 8. Discipline of Research Articles



Co-author citation and co-citation analysis would add another dimension to understanding the dataset as it reflects collaboratory research between the disciplines, but, that, too, is outside the focus of this Study. Figure 9, Discipline of Research Articles, also shows that half of the eighteen articles produced by the Social, Behavioral, and Economics disciplines (revealed in Figure 3) are research-type articles and represent 41% of the total research-type publications.

Table 5 provides a numbers and percentages overview of total publications, research-type publications, and research publications as percent of total publications, by discipline. SBE produced the highest percentage (41%) of the total research publications in the collaboratory literature, but produced relatively

few of the total overall publications, while CS/CCE produced the highest number (34) of overall publications but a significantly lower percentage (21%) of them are research.

Table 5. Frequencies and Percentages of Articles by Discipline

Discipline	Publications	Percent of Total Publications	Research Articles	Percent of Articles that are Research	Percent of Total Research Articles	Research as Percent of Total Publications
CS/CCE	34	40%	7	21%	32%	8%
SBE	18	21%	9	50%	41%	10%
LIS	8	9%	3	38%	14%	3%
Other	26	30%	3	12%	14%	3%
Totals	86		22			

Returning again to Figure 3, Computer Science/Computer Communications Engineering disciplines produced 40% of the total documents, thirty of which focus on Implementation, and two each on Colleagues and Data. Coanalysis with the Article Type shows that seven, or 21% of CS/CCE's articles are Research, and that they represent 32% of the total Research articles and 8% of the total collaborative literature.

The Discipline x Focus Taxonomy revealed that "Other" disciplines produced twenty-six of the eighty-six documents in the collaborative literature, and that eighteen of those articles are about Implementation, six are about Colleagues, and two are about Data. Coanalysis using the Article Type Taxonomy shows that three, or 12% of the "Other" disciplines' publications are Research publications, and that they represent 14% of the Research articles. Other disciplines contribute 3% to the total Research literature of the collaborative.

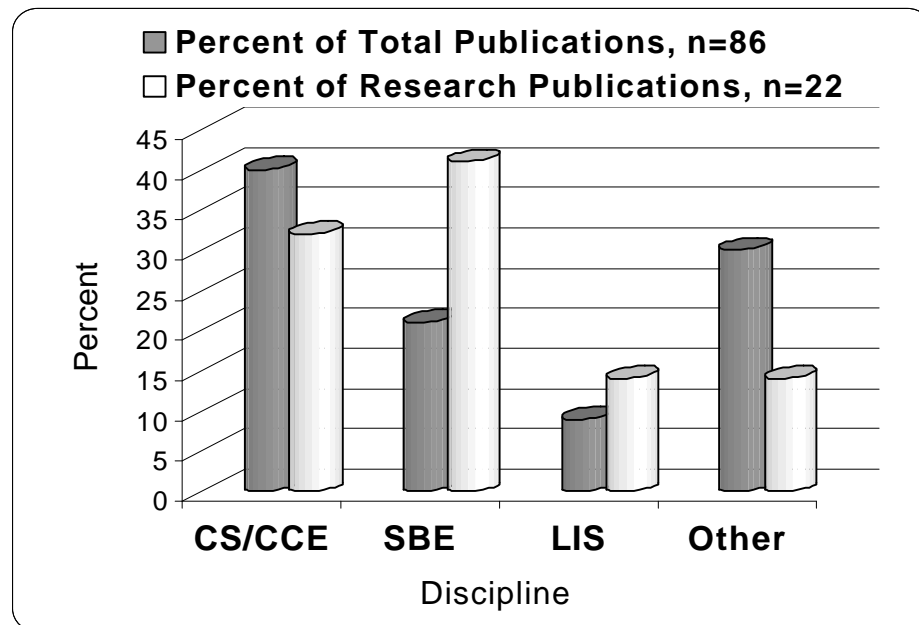
The Discipline x Focus Taxonomy reveals that Library Information Science produced eight of the sixty-eight documents in the collaborative literature, five of those eight publications are about Implementation, two are about Colleagues, and one is about Data. Coanalysis with the Article Type Taxonomy reveals three of LIS's eight articles, or 38%, are research-type articles and that LIS Research contributes 14% to the total research literature, and 3% to the total collaborative literature.

This wearisome deconstruction of the collaborative Literature exceeds Miller's (1956) "Magical Number Seven, Plus or Minus Two" limit of human information processing capacity, and forbids elegant graphing, so an alternative reconstruction is undertaken.

Figure 9 compares percent of disciplinary contributions to total publications (n=86) and percent of disciplinary contribution to research-type publications

(n=22). LIS and SBE each produce proportionally higher percentages of research-type articles in their total publications than do the CS/CCE and Other disciplines.

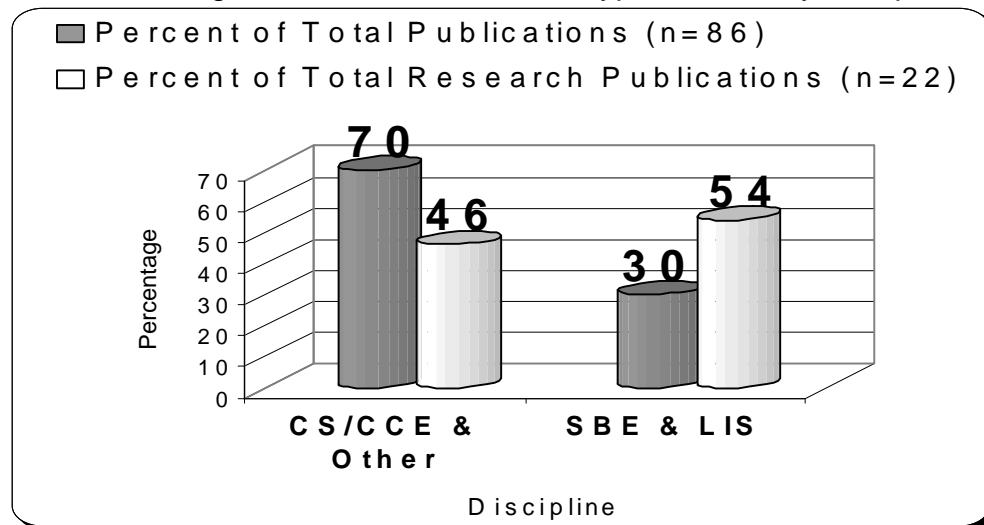
Figure 9. Percent of Total and Research Publications by Discipline



When LIS and SBE are collapsed into the larger category of "Social Sciences," and CS/CCE and Other are collapsed into the larger category of "Hard Sciences" as in Figure 10, an inverse relation between percent of total and

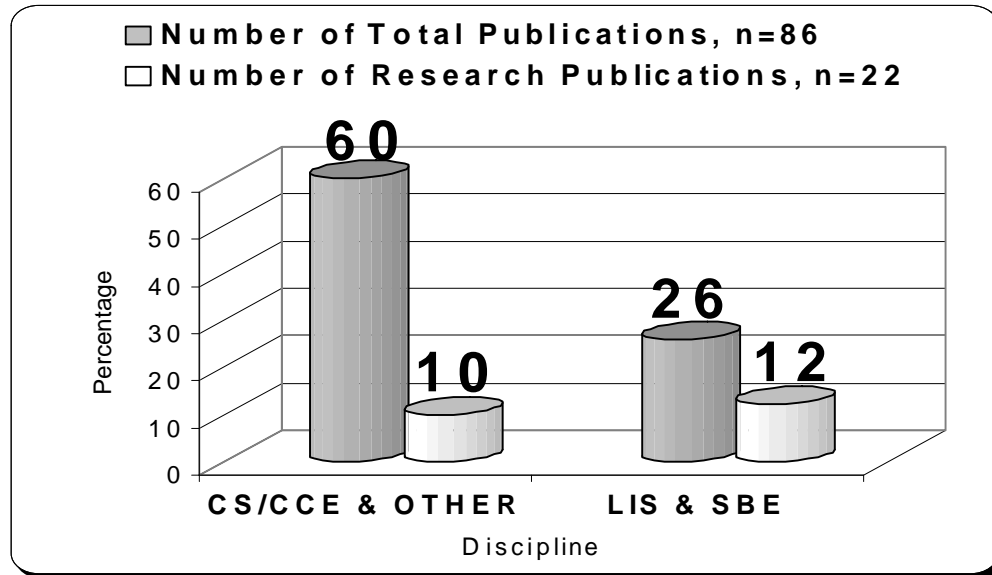
percent of research-type publications is more dramatically revealed.⁴ The social sciences produced more research articles as a percentage of their total publications than did the hard sciences. However, if the percentages are converted to numbers, as in Figure 11, the data show that the hard sciences produced more overall publications than did the social science, but the social sciences produced a relatively equal number of articles that are research.

Figure 10. Percentage of Total and Research-Type Articles, by Discipline



⁴ Although the "Other" Category contains several contributions from disciplines perhaps better classified as "Social Science" than "Hard Science," such as journalism and economics, the contributions of these disciplines was minimal....only one article each. Thus, further reclassification is not deemed necessary.

Figure 11. Number of Total and Research-Type Articles, by Discipline



Conclusion

Wulf (1988) stressed the importance of disciplinarily-balanced research and relative equality of contribution to research aimed at developing the collaboratory, and assumed that research leading to the collaboratory would reflect this balance and equality of contribution. This chapter constructed taxonomies from Wulf's *White Paper* (1988) and Haddow's (1997) document types to analyze the collaboratory literature for disciplinary contribution, research focus, and article types, and to test those assumptions.

The hard sciences and the social sciences made relatively equal, though significantly different contributions to the literature representing the first decade

of the collaboratory. The hard sciences produced a higher overall number of publications, but most of them are Glad Tidings and Testimony and News-Type articles rather than Research-type. The social sciences produced significantly fewer overall publications but a greater percentage of them are Research-type publications. The number of Research-type articles produced by the hard sciences and the social sciences is relatively equal.

This relative equality in number of publications indicates that the scientific community, as a whole, has responded positively to Wulf's interdisciplinary assumption about the environment of the collaboratory. The balance between discipline, focus, and type of publications within the collaboratory literature supports the assumption of relative equality of contribution to, and interdisciplinarity of the collaboratory environment.⁵

In Chapter Four of this study, a taxonomy of research type and topic based on the Lederberg and Uncapher (1989) report is developed, and the collaboratory literature is reanalyzed and coanalyzed from a triangulated taxono-

⁵ An examination of the words interdisciplinary and multidisciplinary is warranted. In this study interdisciplinary is in reference to the environment of the collaboratory, not to the individual or collective articles used to analyze the environment.. Many, if not most, of the articles are disciplinary, in that they reflect the knowledge base and epistemology of a discipline-based scholars. Together, the articles are multidisciplinary, in that they reflect contributions from many individual disciplines. The environment the articles constitute, however, is interdisciplinary, in that it includes the intellectual space "between the disciplines." See Klien (1990) for a full explication of the concept of interdisciplinarity.

bibliometric perspective to further test the assumption of relative equality of contribution to and interdisciplinarity of the collaborative environment. Chapter Five undertakes a qualitative, synoptic analysis of a subset of Theory-Type Research articles (n=22) to lay a foundation for development of an environmental theory of the collaborative.

CHAPTER FOUR

Topic and Approach of Publications

In Chapter Two, William Wulf's (1988) *White Paper*, which sets the philosophical foundation of the collaboratory, and Lederberg and Uncapher's (1989) report, which sets the intellectual foundation of the collaboratory, were presented in detail. In Chapter Three, the collaboratory literature (n=86) is described and analyzed using constructed taxonomies based on Wulf's disciplines needing to contribute to, and the needed focus of research leading toward development of the collaboratory, and on Haddow's (1997) article types. Analysis of the research-type publications (n=22) by discipline and focus revealed that the social sciences and the hard sciences have made relatively equal contributions to the collaboratory's research literature. This supports Wulf's assumption of relatively equal contribution to and inherent interdisciplinary of the collaboratory environment.

Chapter Four continues taxonomic analysis of the collaboratory literature by turning to Lederberg and Uncapher's (1989) *Toward a National Collaboratory: Report of an Invitational Workshop*. Lederberg and Uncapher's report identifies three research topics and three approaches of research required for

development of a national collaboratory. A taxonomy based on Lederberg and Uncapher type and approach is developed, and the collaboratory literature is reanalyzed. A triangulated analysis using the Discipline x Focus, Article Type, and Type x Approach Taxonomies explores a subset of Theory-Type Research publications (n=22), by discipline. Qualitative synoptic analysis of the content of the Theory-Type Research publications is undertaken to balance the triangulated, quantitative, taxonomic analysis and explore for an emergent environmental theory of the collaboratory.

The Lederberg and Uncapher Taxonomy

The Lederberg and Uncapher's report identifies three topics that need to be researched for the collaboratory to be realized:

- Systems Architecture
- Tools and Technologies
- Uses and Testbeds

These topics correspond loosely with Wulf's foci on instrumentation, colleagues, and data. However, Lederberg and Uncapher's topics do not address data as a separate area, and Systems Architecture (networks and networking) and Tools and Technologies (computers and software) are separated into two categories while they are combined in Wulf's Instrumentation category. Also, the

emerging environment of the collaboratory, the testbed, is coupled with uses and user studies, as they might be logically but are not specifically in Wulf's Colleagues category. These slight but significant differences warrant a third taxonomy for analysis of the collaboratory literature and provide an opportunity to validate the findings of relative equality of contribution and interdisciplinarity.

The Lederberg and Uncapher report also identifies three functional approaches of research that needed to be done:

- Design
- Implementation
- Testing

The philosophical approach of Wulf's *White Paper* does not address the need to research the processes of building the collaboratory, as does the Lederberg and Uncapher report with its approaches of research. The attention to more practical matters in the Lederberg and Uncapher report begins building the intellectual foundation for the collaboratory.

A 3x3 taxonomy is constructed based on Lederberg and Uncapher's research topics and approaches, and a third coding of the eighty-six-document retrieval set is performed. During analysis, the literature quickly revealed that a fourth and very distinctive approach to collaboratory research was being taken: theory research. These theory type articles did not fit comfortably into any of the

Lederberg and Uncapher research approaches. Accordingly, a new category, Theory, was added to the matrix and is used for classification of publications concerned with construction of general intellectual models or the development of theory intended to support Implementation, Design, and Testing work. Again, documents might logically belong in multiple categories, but each was logged only once, based on its primary topic and type.

The leftmost column in Table 6 represents Lederberg and Uncapher's topics of research: Systems Architectures, Tools and Technologies, and Uses and Testbeds. The remaining columns represent the three approaches identified as necessary for development of the National Collaboratory: Design, Implementation, and Testing, with the fourth category, Theory, which emerged from the data, added.

Table 6. Topic and Approach of Articles

	Approach				
Topic	DESIGN	IMPLEMENTATION	TESTING	THEORY	TOTAL
System Architecture	9	8	0	9	26
Tools and Technologies	8	11	3	4	26
Uses and Testbeds	5	12	6	11	34
TOTAL	22	31	9	24	86

As an example of coding decisions, a published document representing the development of a multimedia email interface for use in the collaboratory is categorized in the Tools and Technologies/Design cell, while a paper concerned with interpersonal communication theory in the electronic environment of the collaboratory is placed in Uses and Testbeds/Theory cell. Again, each cell was assigned a number, and the number ascribed to each document was transferred to a created field in the study database. The database records were imported into a spreadsheet for analysis and graphing.

There is little difference in the number of publications within each of the topics. Number of publications by topic is graphed in Figure 12. Uses and Testbeds includes thirty-four articles, and Systems Architecture, and Tools and Technologies each include twenty-six.

Figure 12. Topic of Articles

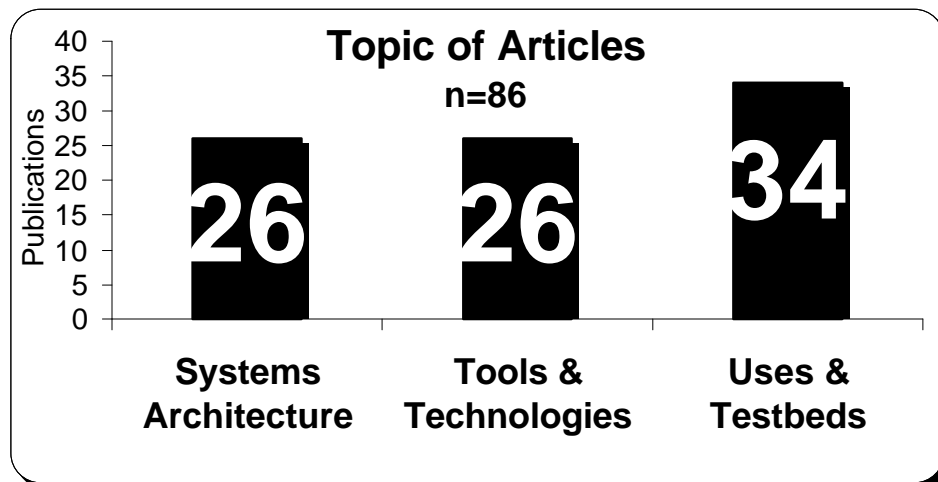


Figure 13 shows that among the eighty-six documents, Implementation-type publications dominate with thirty-one articles, and the added category, Theory, contains the second largest number, twenty-four of the total. The Design-type category includes twenty-two articles, and the Testing-type includes nine. Figure 14 shows the full Topic x Approach taxonomy and reveals a relative equality in dispersion among the categories.

Figure 13. Approach of Articles

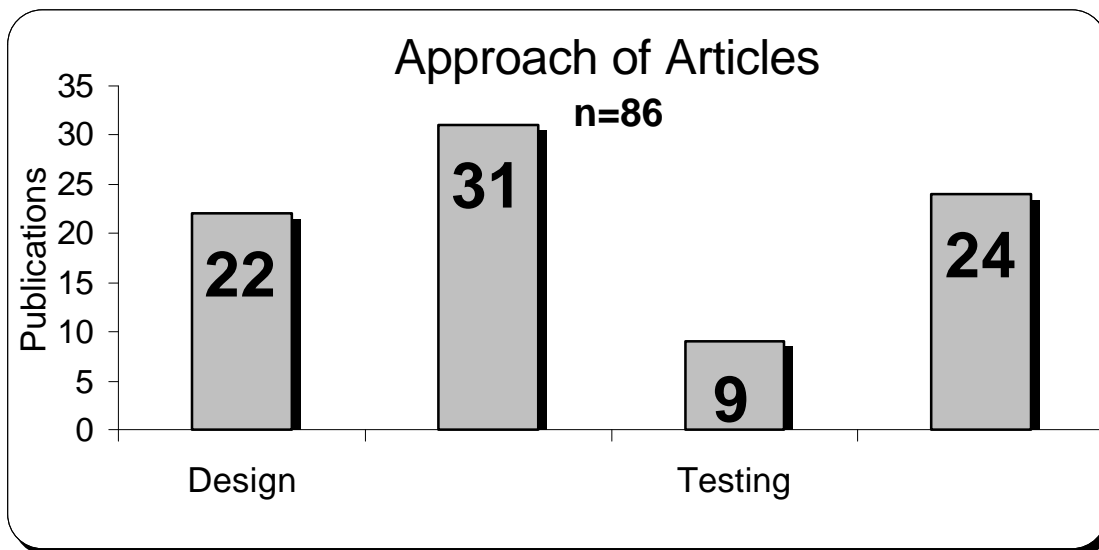
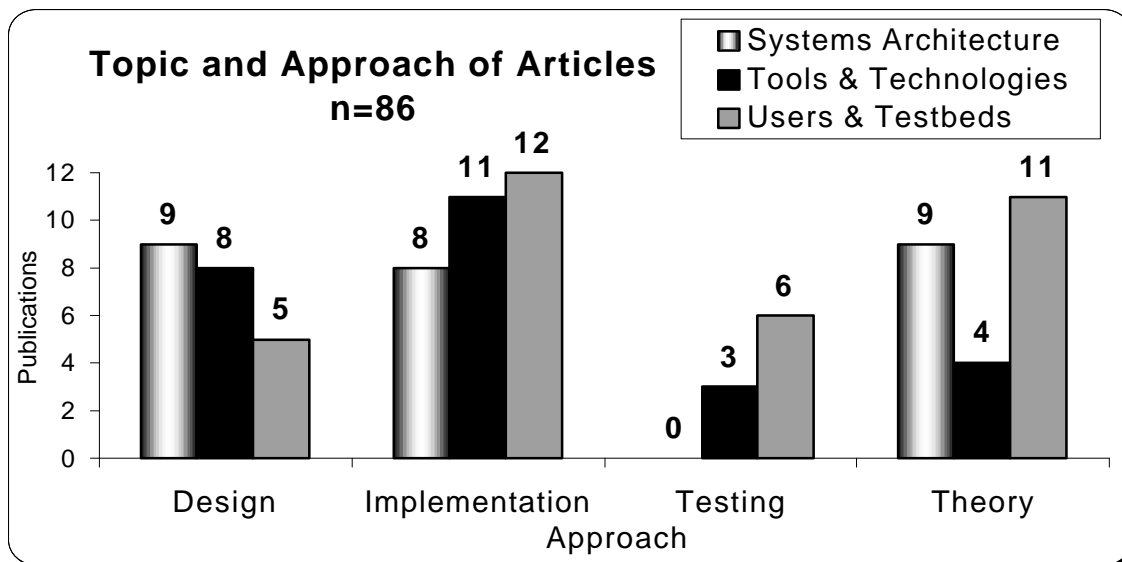


Figure 14. Topic and Approach of Articles



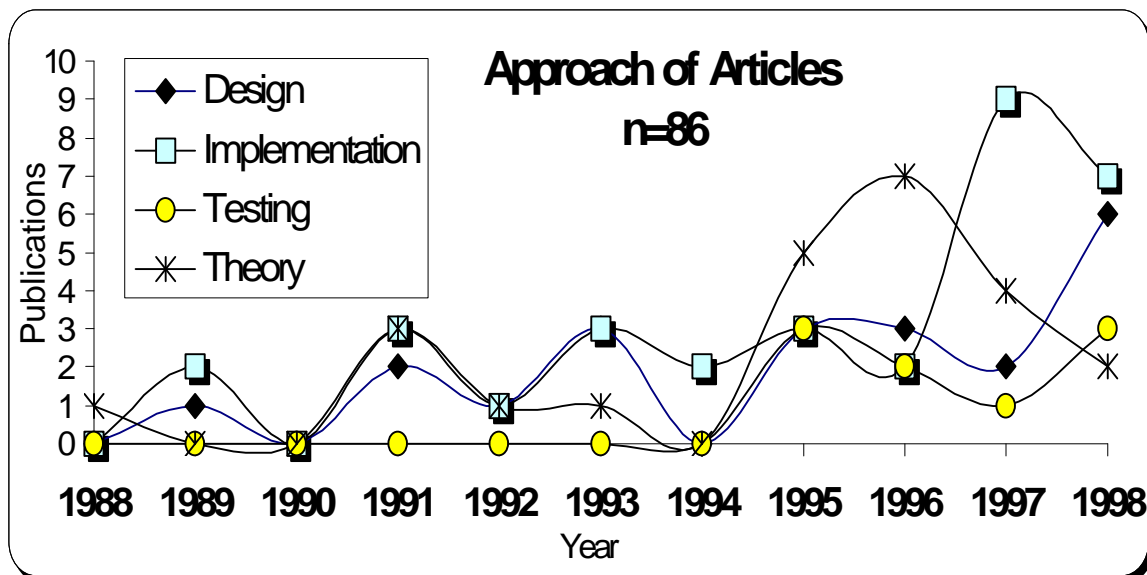
Lederberg and Uncapher's "Cycles"

The Lederberg and Uncapher report suggests that collaboratory research will take a three-fold approach (8), and that the three approaches will repeat themselves in cycles of design, implementation, and testing. To test this, the collaboratory literature was plotted using the Topic x Approach taxonomy, over time.

In Figure 15, each line represents one approach of publications plotted annually. Classic time-series data analysis requires a minimum of fifty datapoints for confidence, so the dataset does not allow the luxury of statistical inference, but early analysis hints that publication by approach does appear to be cyclical.

The widest possible view shows a slight cyclical irregularly beginning in 1994, when theory and then implementation publications make positive surges. Continued monitoring from this perspective would allow inference to such areas as predicting fruitful focus for short and mid-term research agendas. Because the theory category was added during coding of the dataset, and because of the relatively high number of publications placed into it, the theory category was selected for closer analysis.

Figure 15. Approach of Articles Annually



Theory-type Research Publications

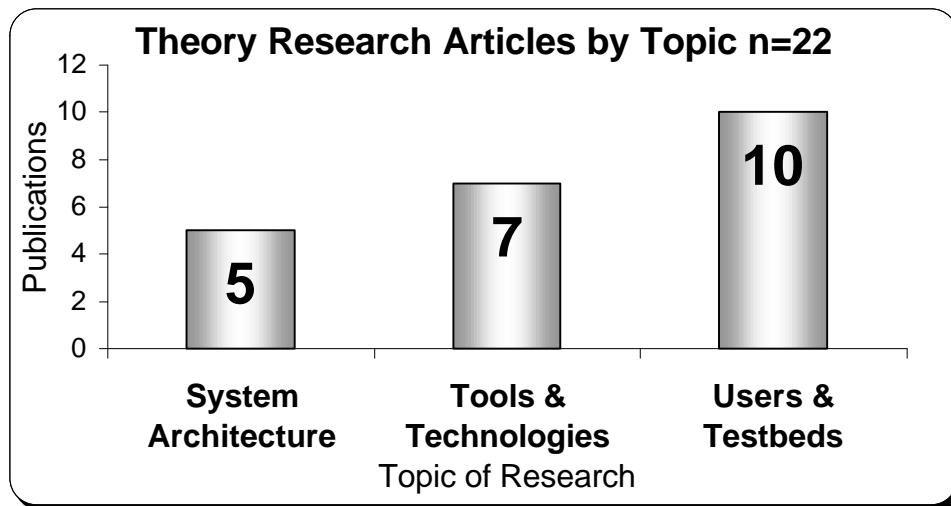
The twenty-four publications identified in the Topic x Approach Taxonomy's Theory category were coanalyzed with the twenty-two documents classified as Research in the Article Type taxonomy developed in the previous chapter. The Article Type taxonomy classes publications as either Glad Tidings and Testimony, Research, or News-Type. The resulting subset is Theory-Type Research publications.

One of the Theory documents not included in both sets is classed as a Glad Tidings and Testimony article because it did not develop or exercise theory, merely speculated about it, and the other is classed as a News-Type article

about collaboratory research generally. Both are removed from the document set, which is reduced to twenty-two publications.

The subset of Theory-Type Research articles is further classified by topics of research and graphed in Figure 16. Ten of the theory-type research publications are classed in the Users and Testbeds topic category. Seven are classed in Tools and Technologies, and five are classed as System Architecture.

Figure 16. Theory-Type Research Articles by Topic

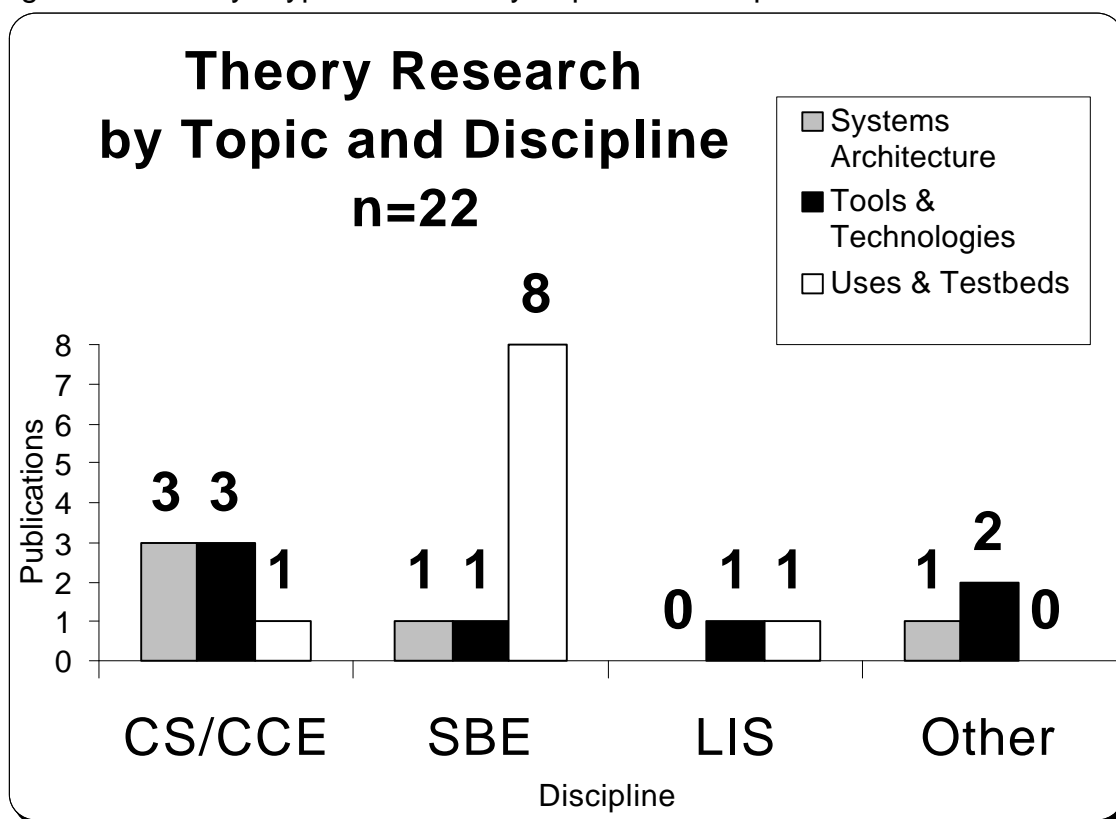


Theory-type Research Articles by Discipline

When the twenty-two Theory-Type Research articles are classed by discipline and plotted in Figure 17, we see that the Social, Behavioral, Economics (SBE)

disciplines contribute the greater overall number as well as the greater percentage of Theory-Type Research articles, and that the Uses and Testbeds Topic produced the greatest number of Theory-Type Research articles.

Figure 17. Theory -Type Research by Topic and Discipline



Social, Behavioral, and Economics disciplines (SBE) contribute ten Theory-Type Research articles: eight about Uses and Testbeds, and one each about Systems Architecture and Tools and Technologies. Computer Science/Computer

Communications Engineering (CS/CCE) contributes seven Theory-Type Research articles: three each in Systems Architecture, and Tools and Technologies, and one in Uses and Testbeds. Library Information Science (LIS) contributes two Theory-Type Research articles: one each in Uses and Testbeds and Tools and Technologies, but none in Systems Architecture. Other disciplines (including physics, biology, botany, chemistry, medicine, journalism and others) contribute three Theory-Type Research articles: one in Systems Architecture, two in Tools and Technologies, and none in Users and Testbeds.

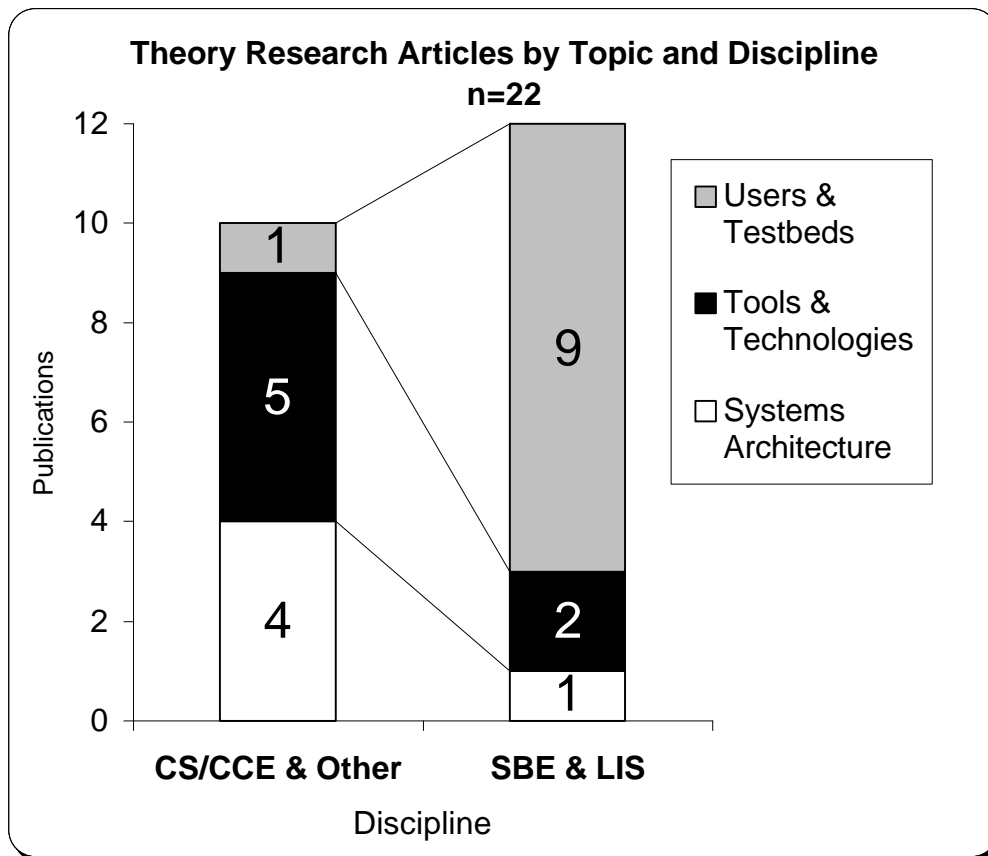
Following the precedent set in Chapter Three, reconstructing the disciplines by collapsing CS/CCE and Other into "Hard Sciences" and SBE and LIS into "Social Sciences," another subtle balance in the contributions to the collaborative literature is revealed, and is graphed in Figure 18. While the hard sciences contribute ten and the social sciences contribute twelve Theory-Type Research articles, the topic of research reported reveals a complementary balance between the disciplines: between Use and Testbed, or human-oriented research, and Systems Architecture and Tools and Technologies, or technology-oriented research.

The hard sciences, or CS/CCE and Other, produced a greater number of Theory-Type Research articles dealing with Systems Architecture and Tools and Technologies while the social sciences, or SBE/LIS, produced an inversely

proportional number of Theory-Type Research publications dealing with User and Testbeds. This finding further confirms as practiced the assumptions of relative equality of contribution to, and interdisciplinarity of, the collaboratory environment as put forth by Wulf and Lederberg and Uncapher in their foundation documents. It also provides the foundation for a statement of an objective reality of the collaboratory: that the environment of the collaboratory, as represented by the published literature, reflects a relatively equal contribution from the disciplines (the body of articles is multidisciplinary), and that the environment of the collaboratory they constitute is inherently interdisciplinary when analyzed by one if not all of Klein's (1990) four strategies for definition of interdisciplinarity:

1. *by example*, to designate what form it assumes;
2. *by motivation*, to explain why it takes place;
3. *by principles of interaction*, to demonstrate the process of how disciplines interact; and
4. *by terminological hierarchy*, to distinguish levels of integration by using specific labels (55).

Figure 18. Theory-Type Research Articles by Topic and Combined Disciplines



Conclusion

Chapter Four examines the collaboratory literature (n=86) and the subset Theory-Type Research publications (n=22) using a taxonomy constructed from the topics and approach of research suggested by Lederberg and Uncapher (1989). Triangulated analysis with the Discipline x Focus taxonomy constructed

from Wulf's *White Paper*, and the Articles Type Taxonomy constructed from Haddow's (1997) publications types in Chapter Three, confirms the relatively equal contribution in number of collaboratory research articles from the hard and soft sciences. A subtle complementarity between approach by discipline within the relative equality of total number of publications, is also revealed, with the soft science disciplines contributing more in the human-centered areas of research (user testbeds) and the hard science disciplines contributing more in the technology-centered areas of research (tools and technologies, systems architecture). Chapter Five examines the twenty-two Theory-Type Research articles qualitatively, and seeks an emergent theory of the environment of the collaboratory.

CHAPTER FIVE

Toward a Theory of the Collaboratory

Chapter Two identifies the philosophical and intellectual assumptions of the relative equality of disciplinary contribution to, and interdisciplinary of, the environment of the developing collaboratory as put forth in Wulf (1988), and Lederberg and Uncapher (1989). Chapter Three verifies these assumptions as practiced, and Chapter Four confirms these practices as principles reflected in collaboratory publications. These findings provide the first statement of an objective reality of the collaboratory: that the collaboratory, as an information environment, is constructed from a relatively equal contribution from the disciplines, as an inherently interdisciplinary environment.

This study now turns to qualitative techniques for a deeper understanding and to explore for an emergent theory of the environment of the collaboratory. The content of the twenty-two Theory-Type Research publications identified in the previous Chapter (and cited with asterisks in Appendix A) are analyzed using constant comparative techniques in the original spirit of Grounded Theory (Glaser and Strauss 1967).

The twenty-two Theory-Type Research publications represent 26% (n=22) of the total collaboratory literature (n=86). These twenty-two articles are Theory-Type Research articles, as distinct from topic-specific but non-theoretical research articles about Systems Architecture, Tools and Technologies, or Uses and Testbeds. The Theory-Type Research articles are further distinct from the two non-research theory-type articles eliminated earlier, and from the Glad Tidings and Testimony and News-Type articles analyzed in Chapter Two.

Systems Architecture Theory-Type Research

Five of the twenty-two Theory-Type Research articles address the topic of Systems Architecture (Wilson 1991, Jaffay et al. 1992, Haga 1996, Huang 1996, Rugelj and Svirgelj 1997). Multiple disciplines contribute to this class of articles (Computer Science, Knowledge Engineering, Medicine, and Psychiatry), and the articles also have an international flavor both in author and journal of publication (United States, Slovenia, Japan, United Kingdom).

Wilson (1991) takes an historical approach to explain that the collaboratory relies on two factors: underlying technologies and group process issues, but contends that these two must be considered together. Jaffay, et. al. (1992) present a conceptual construct of the process of collaboration as one of inclusion into a consistent, coherent, correct whole that is equally viable in the collective

synchronous and individual asynchronous modes. Huang (1996) develops a four layer collaboration model, but contends that the collaboratory can only be grasped in its whole, that changes in one layer reverberate through all levels, and that the ability to adapt and change are of paramount importance. Haga (1996) develops a four layer generic model of collaboration based on participants' activities and (their functions within and commitment to) information flows, and maintains that individual participants and flows have a nonsymbiotic existence within the whole. Rugelj and Svigelu (1997) present a model of the collaboratory based on a medical implementation, and contend that effective communication is the key function, and the selection of instruments should be based on need of the users. The common theme among these System Architecture Theory-Type Research articles is inter- and intra-systems communication, integration, adaptability, and independence supported by individual participation within an indivisible and cohesive whole.

Tools and Technology Theory-Type Research

Seven of the twenty-two Theory-Type Research publications are about Tools and Technologies (Schooler 1991, Rice, More and D'Ambra 1995, Harper and Sellen 1995, Fox and Furmanski 1995, Karamuftuoglu 1997, Citera 1998, Ashton and Levy 1998). Schooler (1991) reviews the history and development of video

teleconferencing and predicts that multimedia teleconferencing has come of age. Rice, More, and D'Ambra (1995) explore Media Richness Theory and identify choice based on situational context as a deciding criteria for the suitability of new media use. Harper and Sellen (1995) explore Media Richness in organizations, and use the cultural inertia argument to explain the lingering preference for paper, concluding that information that requires judgment in its production is less easily shared than information that does not require judgment; and that social interaction is not as crucial to the sharing of objective information as it is to sharing of interpreted information. Fox and Furmanski (1995) explore the future of the Web and conclude that the vision of the Web can only be achieved with truly open and pervasive technologies. Karmanuftuoglu (1997) argues for the semiotic approach to information storage and retrieval, and uses choice as a foundation on which to resolve the conflicting acts of detonation (description) and prescriptions (performatives) as forms of computer language for development of information systems. Citera (1998) looks at the impact of communication media on influence and decision quality, concluding that dominant personality types maintain levels of influence across media, while less dominant types gain influence when using technology for communication. Ashton and Levey (1998) explore the impact of network learner support and the new roles for service and staff with responsibility for promoting skilled use of electronic information

resources, concluding that new roles based on the sharing principle are developing. The overarching concept of these seven Tools and Technologies Theory-Type Research articles is equalization in communication via media richness empowered by choice, power, openness, and sharing.

Uses and Testbed Theory-Type Research

The ten articles in the Uses and Testbeds Theory-Type Research retrieval set (Kydd and Ferry 1991, Mantovani 1995, Barua et. al. 1995, 1996-1997, Robbin 1995, Travica 1995, Mitchell and Singh 1996, Andersson and Roonnberg 1996, Glasner 1996, Swanson et. al. 1997) include all eight of the articles produced by the Social, Behavioral, and Economic (SBE) disciplines. Kydd and Ferry (1991) provide an integration of the literature of Computer Supported Cooperative Work and Media Richness Theory. They use behavioral theory to argue that information processing occurs during group work for two reasons: to reduce uncertainty, and to reduce equivocality; they contend that matching the situation with the appropriate tool is a critical implementation decision.

Mantovani (1995) explores virtual reality from the social psychology perspective as being a communication environment involving consensual hallucination, fiction and possible selves, and integrates a theoretical framework centered on self identity processes. Barua, Chellappa, and Whinston provide two

of the publications (1995, 1997). The 1995 document recounts the experience of creating a collaboratory for the business environment and explores complementarity theory as a way to evaluate value that users derive from the system. The 1997 publication recounts the design and development of an Internet and intranet-based collaboratory, analyzes the needs of three geographically dispersed electronic communities, and concludes that a shift of technology platforms from proprietary to open environments is necessary. Robbin (1995) recounts the construction of a collaboratory based on delivery of the federal government's massive *Survey of Income and Program Participation* (SIPP). She explores the use of distributed database software for delivery of partial information from a large dataset using internetworked technologies, and analyzes the success of the project based on the NRC (1993) *Report* (both Robbin and NRC are discussed further in Chapter Six). She concludes that focus on communication flows, the data-to-information-to-knowledge process, and alleviating administrative bottlenecks are necessary for progress. Travica (1995) explores the accounting industry for the culture of collaboration, concluding that traditional cultural propositions largely hold online, and that it is possible to create a profile of the typical collaborating professional.

Mitchell and Singh (1996) explore the survival of businesses that use collaborative relationships, and conclude that those that collaborate are less

likely to shut down than businesses that follow independent approaches when the environment changes gradually, but that businesses using collaborative relationships are susceptible to takeover. They also find that businesses that collaborate and suffer sudden environmental shock in the area of the collaboration are more likely to shut down than collaborating businesses that suffer sudden environmental shock outside the area of collaboration, concluding that inter-firm collaboration is usually beneficial.

Andersson and Ronnberg (1996) find a difference among individuals who collaborate during two types of memory tasks (semantic memory tasks and episodic memory tasks). Friends as opposed to non-friends reduce the negative impact of collaboration and the "free-ride" or "hide-in-a-crowd" phenomenon frequently associated with collaboration has no impact on the success of collaborators in either type memory tasks. Glasner (1996) uses social theory based on Merton's view of science to investigate the changing culture of science based on the collaboratory, and finds that the collaboratory model has no real impact on accepted views of the culture of science. Swanson, Bailey, and Miller (1997) investigate how entropy as a measure of system disorganization occurs and is measured in physical, biological, and social systems from the perspective of exchange of money-information markers. The underlying theoretical theme among these ten articles is fair exchange, sharing, and commonality, with

maintenance of strong individuality within the collective collaborative environment.

Theory Development

Synoptic statements from the abstracted content of the twenty-two Theory-Type Research articles are:

- Integration and adaptability is necessary and good.
- Change, choice, and personal power are requisite.
- Consensus, sharing, and exchange are positive and practiced.
- Individuality and collectivity are distinctly and respectfully maintained.

Taken as a whole, these qualitative findings point first to the remarkable absence of even the most subtle trace of masculine (hierarchical or patriarchal) social behaviors, among them individualism, dominance, competition, confrontation, mastery, aggression, advantage, etc. (Crimshaw 1986). Second, they resonate with an underlying harmony antithetical to traditional scientific and technological practices.

Without mentioning gender, Wulf (1988) and Lederberg and Uncapher (1989) project that the collaborative would be antithetical to the traditional practice of (male dominated) science. They expected competition to be replaced with cooperation. They expected individual work to be replaced by group work. They

expected individual rewards to be replaced by group rewards. They expected traditional administration and control mechanism and notions of ownership to be supplanted by new, decentralized systems based on sharing. They also projected it would be difficult to convince practicing scientists to accept these new notions about how science might be done.

McLuhan (1963) warned that the "medium is the message," and that we are entering a tribally-oriented "global village," in which electric technology is an "extensions of the self," leading toward a "world soul" or universal human nervous system. Teilhard de Chardin (1975) observed that incorporation of technology is part of natural, uninterrupted human evolution leading toward an etherized human consciousness and a unification of the nervous system. If technology is about the process of knowledge, about *how* we know, rather than the product of knowledge, or *what* we know, as it was in the mechanical age (Castells 1996), the early ideals, philosophies, and theories of the technologically-determined collaboratory suggest that these fundamental intellectual changes have been incorporated in the theoretical design of the emerging collaboratory environment.

Dorman (1998) finds that girls and boys interact with technology differently. Boys view technology as a way to extend their power and play games based on competition and contest, finding the workings of technology itself as enthralling

as the uses of technology, whereas girls use technology as a way to connect with people and solve real life problems. In light of the fact that technology and science are male dominated professions (NSF 1999, Raber 1998, Clegg 1999, Weinman 1999), and the collaboratory is a male construct, this raises the question of feminism.

It is dangerous to draw dichotomous generalities about gender, and equally difficult to describe a fundamental feminism (Crimshaw 1986). Nevertheless, collaboration is a social practice at which women, generally, excel over men (Clinchy 1985, Roschelle 1992) and the collaboratory is an environment purposively constructed to foster collaboration. Certainly, engendering female attributes does not make the collaboratory feminist; that would be a sort of "this is not a duck" definition of a goose. It does, however, suggest a purposively de-gendered masculine environment, or an ungenderedness, which in many circles is remarkably feminist (Haraway 1985).

The qualitative statements distilled in this chapter are admittedly synoptic and preliminary, but nevertheless suggest that the collaboratory is at once protective of the individual, yet highly conducive to the collective. This reflects feminism's concern with the relationship between autonomy and dependence; between responsibilities to others and needs for self-realization or self-affirmation (Crimshaw 1986, 260). But, pending substantive ecological or sociological

studies of the people, practices, and values of humans working in the functioning collaboratory, we are limited here to considering that part of the environment which is determined by technology, and specifically, how that environment is reflected in the library literature.

Perhaps the collaboratory, which is built of traditionally male tools but purposively designed to accommodate predominantly female behaviors, is a large step toward the unification of the human nervous system. Perhaps American scientists have built an ideologically utopian (Mannheim 1936) feminist environment, a sort of intellectually equalized opportunity to make new realities that reflect a holistic human way of knowing. Unlike the public Internet, which is open to all who would connect, the collaboratory is a closed environment of highly trained and educated intellectuals and visionaries, and so in many ways is protected by protocol from the overt and covert sexual exploitation and gender objectification behaviors that proliferate online. Perhaps the American scientist, by accommodating women's ways of knowing (Goldberger 1996), has in the process set himself intellectually free from the self-imposed limits of men's ways of knowing. It is too early to draw such broad conclusions, which is not the point of this study anyway. It is not, however, too early to consider the ungendered, harmonious nature of the collaboratory as informing the sociological and ecological studies that are required to sustain and advance the collaboratory.

Joining Haraway, feminist scholars Herring (1994) and Mallon (1998), among others, find fertile ground in technologically-enabled internetworked environments for the propagation of women's ways of knowing. They see great potential, but also the need for constant vigilance lest old constructs propagate unchecked in these new environments. Kelley (1963, 161) identified three conditions favorable to the formation of new constructs, including use of fresh elements (specifically setting), experimentation, and availability of validating data, all of which are available in the collaboratory. Wilson (1999) calls for a consilience or unity of knowledge between the hard and social sciences; and McLuhan projects that technological media will require replacing the traditional human classification imperative with pattern recognition behaviors. Castells (1996, 32) sees technological uptake as an evolutionary process moving from the accumulation of knowledge to the application of knowledge. Clearly, we are in the midst of great change.

As a connective, problem-solving scientific environment, the collaboratory seems remarkably hospitable to the female ways of using technology. Certainly, nothing in this study so far finds the collaboratory inherently antagonistic to feminism. The collaboratory emerges from the male-dominated worlds of technology and science as an ideologically ungendered, harmonious environment built from relatively equal contributions from the disciplines, as an

inherently interdisciplinary, intellectual environment. Since the "nucleus of the scientific method is the rejection of certain propositions in favor of others in strict conformity to fact-based logic" (Wilson 1999, 264), the theory that the collaboratory is a relatively equal, interdisciplinary, ungendered information environment is put forth, and this study proceeds seeking rejection of that proposition.

As the evolution of the collaboratory passes from the implementation, and tools and technologies stages, and begins to concentrate on productive human uses and designs, focus shifts to the actual (*vis* theoretical) environment, and to the population, or the inhabitants of the new environment, and eventually to the intersubjective "rules of the road" for the collaboratory and the skills collaboratory pioneers value in prospective participants. These last two issues are considered in Phase Three of this study. But first, a subjective, experiential understanding of the functioning collaboratory is necessary. The subjective reality presented in the Phase Two of this study is achieved through prolonged immersion in the online environment of the collaboratory. It serves not only as a synoptic introduction for prospective participants, but tests experientially the validity of the objective ideals, philosophies, and theories the library literature provides. For Phase Two of this study, the collaboratory is not only an object for

the conduct of science, but an environment subject to scholarly investigation: the collaboratory becomes the scientific sample.

Conclusions

Quantitative analysis of the collaboratory literature using taxonomies constructed from Wulf (1988), Haddow (1997), and Lederberg and Uncapher (1989) confirms as practiced principles the assumptions of relative equality of contribution to, and inherent interdisciplinarity of the information environment of the collaboratory. The hard sciences and the social science have made relatively equal contributions to the collaboratory literature, and publication patterns as they constitute a single information environment, are inherently interdisciplinary (Klein 1990, 55). The differences in the contributions of the disciplines in topic, type, and approach of publication equalize within the numbers and percentages of total publications.

Synoptic description of twenty-two Theory-Type Research articles reveals four theoretical themes: that integration and adaptability are necessary; that change, choice, and personal power are requisite; that consensus, sharing and exchange are positive, and that individuality and commonality are maintained within the collaboratory. The overarching commonality among these four distilled theoretical themes: equality, choice, sharing, and consensus, are, by virtue of the

absence of traditional male scientific philosophies, ungendered, and by their very nature, harmonious. An emergent theory of the collaboratory as a harmonious, ungendered, intellectual, information environment is put forth.

CONCLUSION OF PHASE ONE

Phase One of this study, comprised of Chapters Two through Five, constructs an objective reality of the collaboratory based on the holdings of the library. Published collaboratory research reflects the relative equality of contribution to, and inherent interdisciplinarity of, the collaboratory environment. Chapter Two explores two of the collaboratory's three key documents, neither of which was published or widely distributed, but both of which are frequently cited. Wulf's (1988) *White Paper* provides the philosophical foundation of the collaboratory and identifies the disciplines that need to participate in and the focus of needed research for the collaboratory to develop. Lederberg and Uncapher's (1989) report provides the intellectual foundation for the collaboratory and identifies the topics and approaches of research needed. The collaboratory literature (n=86) is defined as those publications made available through the intermediation of the library, retrieved using the search term "collaboratory," and accessed during skilled library research. The collaboratory literature represents only those documents highly pertinent to the collaboratory, is recognized to exclude a substantial relevant literature, and does not represent the total literature of either criterion. The literature spans the first ten years of the collaboratory (1988-1998),

and is recognized of insufficient number to support rigorous statistical analysis. Therefore, taxonomic classifications and descriptive statistical techniques using counts, frequencies, trend analysis, and various coanalyses are used for taxono-bibliometric analysis of the literature.

In Chapter Three, two taxonomies are constructed and analysis of the documentary evidence of the collaboratory is undertaken. The first taxonomy is based on Wulf's (1988) research disciplines and focus, and the second on Haddow's (1997) article types. Chapter Three identifies and confirms the assumption of relative equality of contribution to, and interdisciplinarity of the collaboratory environment as set forth in Wulf's philosophical paper.

In Chapter Four, a taxonomy based on Lederberg and Uncapher's research topics and approaches is constructed and used individually, in tandem, and in triangulation with the Wulf and the Haddow taxonomies to reanalyze the collaboratory literature. Chapter Four confirms the relative equality of contribution to, and inherent interdisciplinarity of, the collaboratory environment, and identifies twenty-two Theory-Type Research articles for closer analysis.

Chapter Five relies on synoptic description to probe the twenty-two Theory-Type Research articles for an emergent theory of the collaboratory environment. Common themes of interdependence, sharing, fair and equal exchange, openness, and preservation of the individual within the collaboratory environment are identified, and a theory that the collaboratory is an harmonious, ungendered, intellectual information environment is put forth.

Phase Two of this study creates a subjective reality of the collaboratory during prolonged immersion in the online environment, develops criteria for inclusion as a collaboratory, and presents four descriptive studies. Phase Three creates an intersubjective reality of the collaboratory via a Delphi among collaboratory pioneers to determine the "rules of the road" for the collaboratory and identify skills collaboratory pioneers value in prospective participants. The study concludes with a philosophical intertwining of the objective, subjective, and intersubjective realities of the collaboratory and suggests areas of additional research.