Chapter 2 --- Models, Research Process, and Collaboration: A Review of Literature.

This dissertation proposes a new way of conducting long-term, large-scale social science research. In order to understand both the present and proposed social systems of research, we need to examine many issues. These issues include the nature of the research process itself, the nature of the scholarly establishment, the supporting technology that served in the past and the technologies that are necessary to support research in the future. While the success of large-scale long-term research in the physical sciences is quite clear, social science research is dominated by solo or small-scale efforts. We need to know why; and when we do we may see how new research efforts can be conducted on a scale that matches our ignorance of social processes and problems.

Figure 2.0 below shows some of the shortcomings of current research methods. This is to be contrasted with Figure 2.5 that shows the conceptual framework of this dissertation.

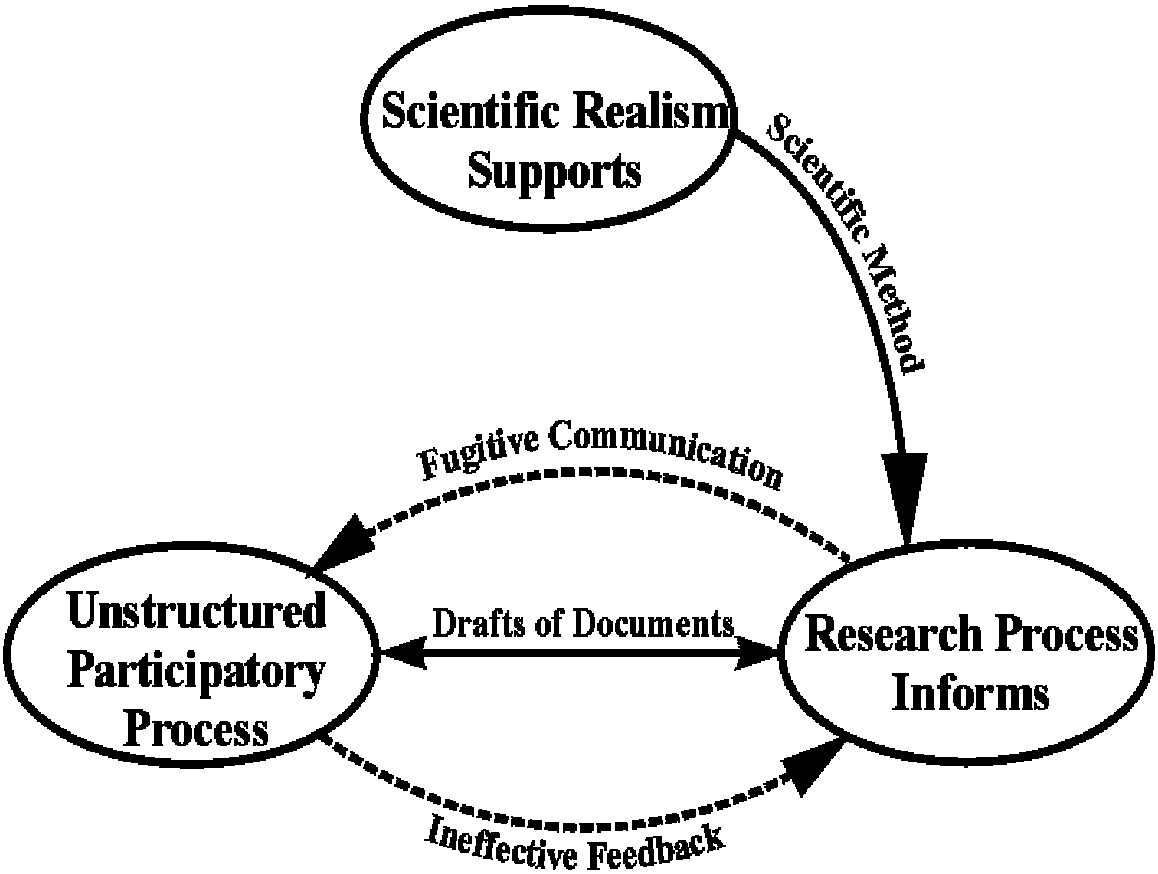


Figure I Existing Collaborative Environment

We start with the selection of scientific realism as an underlying philosophy. The research process is currently driven by the team's understanding of the scientific method and by the customs of their discipline. The participatory process, so necessary in large-scale research, is currently typically unstructured, dominated by unrecorded dialog between the participants -- often on a dyadic basis rather than on a team basis. Participation is now dominated by the sharing of document drafts in a process where

criticism is not shared, but is generally channeled to the team leader. This concentration of power is antithetical to current concepts of social participation.

We hope to remediate many of the problems in current research environments with a new process that again begins with the selection of scientific realism to guide our research. A methodological approach to research will guide the research team from its conception to maturity; we select a methodology that resonates with realism. We end with a proposed conceptual framework that allows us to relate the interacting elements that comprise a research process and technological environment suited to large-scale long-term social scientific research collaboration.

2.1 Scientific Realism and Models

> *"There are only real things and the real ways they behave. And these are represented by models, models constructed with the aid of all the knowledge and techniques and tricks and devices we have."* --- Nancy Cartwright[1]

> *"The Fact is the basis, the foundation; Imagination, the building material; the Hypothesis, the ground plan to be tested; Truth or Reality, the building."*
> ---J.H. van't Hoff (1852-1911) Chemist, Musician, Mathematician, Positivist

A realist believes in an objective existence of the objects found in the research process. The scientific realist might be one who is persuaded by the arguments found in *Realism Rescued: How Scientific Progress is Possible*[2]. The principal argument in that work is that theory is actually embodied in a suite of three models. These models are the source, descriptive, and the explanatory models. The source model is an unexpressed model of reality from which the descriptive and explanatory models are derived. The source model is the union of the tacit knowledge of the team members, the literature that has been identified by the team, and the intellectual content of the current received wisdom. The model may include myth, ideology, and dogma[3]. The descriptive model is a representation of the objects and relationships between the objects of the issue domain.

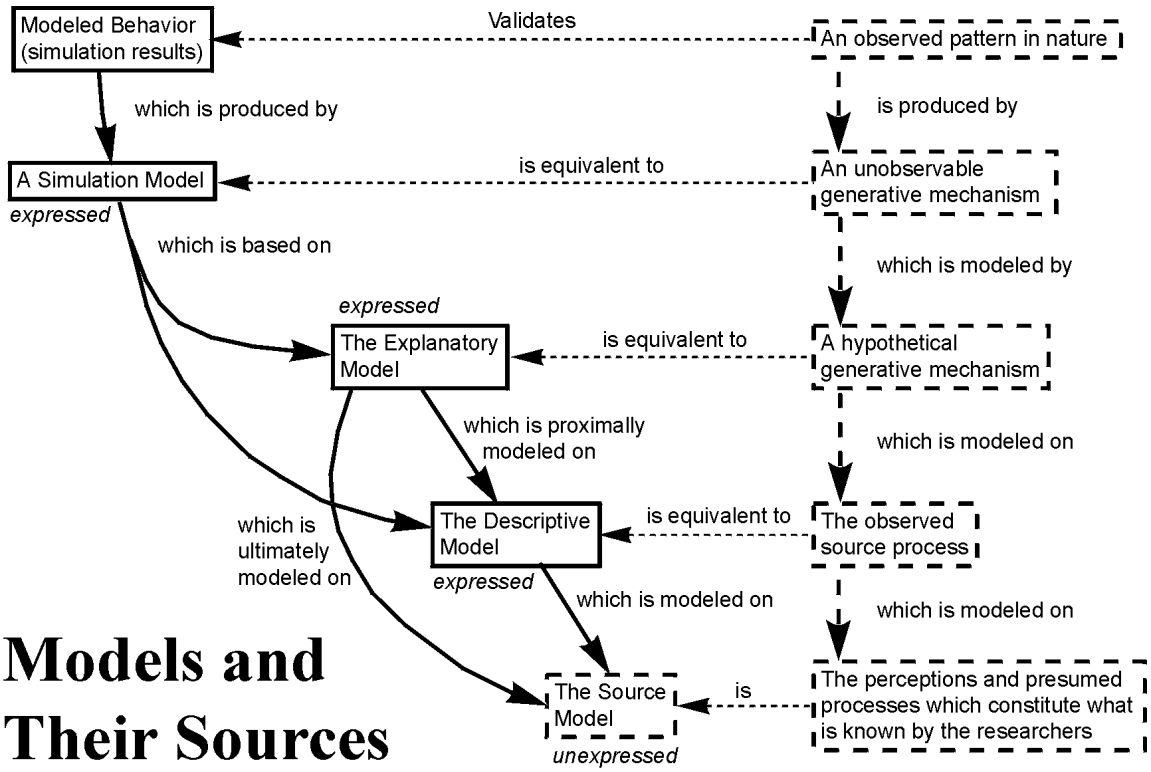
The explanatory model is a representation of the causal power of objects and the mechanisms driving the processes[4].

I have extended the philosophical argument to include a simulation model. This model takes findings from the descriptive and explanatory models and creates an environment that can be used to investigate the effects of change in the system. The simulation model, if provided the state conditions existing in reality at a past time, and the inputs subsequent to that time will produce a description of the system state at the end of the input actions. If the output of the simulation model consistently produces results practically resembling the actual behavior of the system, then the entire chain of modeling from source to simulation is validated by correspondence[5].

The support that scientific realism provides to research is a rationale for modeling practice. The descriptive model is an organized representation of all the information acquired and developed about the objects and their relationships. The explanatory model provides information of an operationalizational nature to the simulation model: qualitative and quantitative parameters. Both the descriptive and explanatory models may be represented by diagrams and supporting text derived from existing genres, especially from software engineering and artificial intelligence[6,7,8,9,10]. Simulation modeling requires much more formality in modeling and calls for the rigor of Unified Modeling Language
(UML)[11,12], Object Role Modeling (ORM)[13] or other formal methods. The simulation model is a computer program represented by algorithms, which are in turn supported by a variety of graphic models and textual documentation. The model representation, as a genre, is discussed at length below (see §3.2.4.2).

We will see that the Research Process, described below in Figure I, is composed of three interdependent domains that are supported by the modeling rationale of scientific realism. The research work within the substantive domain, the observed world, is represented in

the descriptive model. The Explanatory model represents the relationships among elements of the conceptual domain, the causally driven world. The experiments undertaken in the methodological domain, the manipulable world, are reiterated in the Source (simulation) model.



(Harré, 1970, 1978; Aronson, Harre and Way 1995; and Hendricksen)

Figure II Models and Their Sources

2.1.1 The Source Model

The descriptive and explanatory models are based on the team's source model. This source model is simply all the knowledge the team has about the issue domain, the perceptions and presumed processes that constitute what is known by the researchers. It is the union of the tacit knowledge of all research team members, the scientific literature known to members, and any new knowledge produced by the team. Note that tacit

knowledge may contain knowledge that is in fact not true[14]: knowledge that is false, myth, or error. Philosophical myths and dogma may also be a part of each member's tacit knowledge[15]. To a large extent, the source model is the union of all the team member's mental models. Mental models are cognitive categories, beliefs about the world, held by individuals[16]. The source model is unexpressed; but as it becomes expressed its objects and its processes become part of the descriptive model. Like any other model, the source model changes with new knowledge; errors are excised and new knowledge, beliefs and conjectures are incorporated.

Tacit knowledge exists on a continuum of codification of the mental model through culture and routines into expressed knowledge. The qualities of codifiability and teachability influence the rate of knowledge transfer[17]. Moving tacit knowledge to more explicit codified forms is one of the principal jobs of the knowledge network known in this research as the Research Web. The means of transforming knowledge is communication, and communication of tacit knowledge is dependent on the strength of social ties: strong ties are needed to transfer tacit knowledge, while explicit knowledge is easily transferred through weak ties[18].

Since the source model is unexpressed, it cannot be criticized. As elements of the source model become transformed to explicit knowledge and expressed in other models, the other members can subject the elements to critical review. The Research Web is designed to support this criticism.

2.1.2 The Descriptive Model

The Descriptive Model (DM) is the repository for knowledge related to the observed reality of the issue domain. A description of every type of tangible entity (object) that exists within the issue domain has a place in the DM. Every process that is observed to operate in the issue domain is described in temporal, spatial and/or social terms.

Processes relate objects within the issue domain and relate parts of the issue domain to reality outside the issue domain.

As the elements of the source model are expressed they become part of the descriptive model. Objects are identified and refined by extension and criticism. Relationships between objects are expressed and entered into diagrams that organize the knowledge about the objects. Observed processes are described, showing the actual operation of some of the activities within the issue domain. The Research Web, through hypertextual organization, takes its shape from the hierarchical organization of the objects into natural types. Process models relate all the objects and show how they operate in time and space and perhaps socially as well.

The Descriptive Model is created, extended, criticized and refined by the entire team, but most especially by those members who work in the Substantive Domain of the Research Process (see §2.3.2.1, below). As time passes, regardless of the stage of completeness, the Descriptive Model is continually revised and those changes cause ripples to perturb the other models of the system. The DM is the model that the Explanatory Model is modeled on.

2.1.3 The Explanatory Model

The Explanatory Model (EM) is inextricably paired with the Descriptive Model[19]. While the DM seeks to represent the issue domain, the EM seeks to provide causal mechanisms driving the issue domain. The EM is built on the elements, objects and processes, from the descriptive model and the set of hypotheses that emerges from the research. The EM provides a model of a hypothetical generative mechanism for the observations of the DM[20]. For each observed process from the DM there will ideally be a set of explanations based on testable hypothesis. The testing is carried out by experiment to corroborate the developing theory.

The construction of the EM is the principal responsibility of team members working in the Conceptual Domain of the Research Process (see §2.3.2.2 below). The EM becomes the basis for hypothesis generation. As findings accumulate from experiment and research, the EM becomes more mature and comprehensive and eventually becomes a definitive representation of theory.

2.1.4 The Simulation Model

The simulation model (SM) is derived from the DM and the EM. The SM is the model driving the work of those researchers operating in the Methodological Domain of the Research Process (see §2.3.2.3 below). The SM can be used to develop describe hypotheses and develop experiments. In this way the SM provides results for the further refinement of the DM and EM. The SM will be implemented late in the research process, in stage 3 (see §2.3.3.3 below) or the corroborating phase. In many research projects the SM will not be fully implemented due to both time and expense. The simulation model will be a very major undertaking which can really only be justified in long-term, large-scale research projects -- just the project configuration proposed for the Research Web.

Of great importance to the development of the Simulation Model is another model, called the Auxiliary Model, which explains the operationalization of the elements of the SM from their origins in the DM and EM. The auxiliary model concept was proposed by Blalock[21]. A carefully designed auxiliary model insures that measurements of a given attribute of a DM variable are properly represented in the SM[22]. Blalock[23] later suggested that auxiliary models are necessary "whenever our theories are about postulated properties (such as attitudes, values, and abilities) that are only indirectly inferred on the basis of directly observed behaviors." He explores four specific situations calling for careful use of auxiliary models, including very common situations such as the need to use data from two or more levels of analysis, and situations where the measured variable is intrinsically fuzzy or boundaries are rather arbitrary. The auxiliary model is

also necessary to operationalize variables whenever an experiment is designed to explain any phenomenon[24].

The hypertextual organization of the RW is essential to demonstration of the operationalization rationale of every element of the simulation model. Since the auxiliary model is diagrammatic and textual, it may be easily criticized. In short, the auxiliary model provides validation of the SM, under the assumption that the DM and EM are correct.

> *...it seems wise to develop a hypercritical stance that brings many ... hidden assumptions into the open as possible, and then to examine very carefully their implications.* ... H.M. Blalock[25]

In practice the auxiliary model will not exist as an entity, but will be expressed as augmentations to the DM. The nature of the augmentation is to explain how the attributes of objects are measured. The auxiliary model can be thought of as the argumentation that validates every operationalization decision in the SM. Each of these arguments can be independently criticized and refined. With multiple operational definitions of objects multiple versions of the SM may be implemented each with different results to discuss.

The simulation model is based not only in scientific realism, but in general systems theory[26] as well. General system theory considers hierarchy to be fundamental in all systems[27] just as scientific realism bases its models on a hierarchy of natural kinds. This philosophical conjunction supports the existence of a hierarchy of models at different scales or levels of abstractions. These nested hierarchical models would seem to fulfill Morrill's[28] wish for methodology that would support "several levels of explanation for the same phenomena, depending on the question asked."

2.1.5 Grounding of the Use of Models

The models, when well developed, provide support for processes fundamental to the scientific investigation. Charles Pierce argued that scientists engage in syllogistic reasoning in their work, and this results in a 'logic of discovery' of four steps[29]:

1. observation of an anomaly.
2. abduction of hypotheses for the purposes of explaining the anomaly.
3. inductive testing of the hypotheses in experiments.
4. deductive confirmation that the selected hypothesis predicts the original anomaly.

In a well-developed model, an anomalous qualitative or narrative datum will likely be noted and perhaps critically discussed. Certainly numerical data may be subjected to statistical tests that will identify outliers and other instances, which beg explanation. In the development of explanatory models these data represent problems that need to be disposed of, as they defy the current explanations.

The means of explaining anomalies begins with a process of abduction. Abduction is not well understood philosophically, but is well enough known to be identified as a practical tool. Shank[30] associates abduction with 'ground-state' or ordinary thinking. Abduction is applied any time we find ourselves in an ordinary situation and know what to expect[31]. Six modes of abductive reasoning[32,33] are identified as:

a) reasoning to the omen (or hunch)
b) reasoning to the clue
c) reasoning to the metaphor or analogy
d) reasoning to the symptom
e) reasoning to the pattern
f) reasoning to the explanation

In the research environment, omens, hunches, clues, metaphors, analogies, symptoms, patterns and explanations are supplied in abundance through data, observations, ideas,

proposed explanations, etc. The primary repository of these data are: the Descriptive Model; the RW essays, data stores and e-mail archives; and, to a limited extent, the Explanatory Model. Hypotheses developed through abduction to explain anomalies may be presented in documents that may be criticized.

Inductive reasoning is used to examine the formalized and operationalized hypothesis through experimentation. The results of the experiment will then be used as evidence in the Explanatory Model. These explanations should also then be linked from the observation of anomalies that started the abductive examination leading to the hypothesis.

Deductive reasoning is seldom employable in the social sciences. Deduction is the examination of truth based on axioms and rules of logic, and rules and axioms are absent or weak in most social sciences; thus formal deduction is not seen there. In the research environment, however, we do have axioms and rules. Due to the impossibility of knowing everything about an open social system, we build the Simulation Model to summarize and operationalize our knowledge about the system. The Simulation Model must be realized as a computer program, and computer programs operate with rules and axioms.

So, to the extent that our rules (logic and synthesized explanatory processes) and axioms (observations) are true, we can deduce or predict. Now, if we submit a scenario to the Simulation Model, and the Simulation Model predicts an outcome that is compatible with the expected outcome, then we have a weak confirmation of the models. That weak confirmation is, in effect, an observation made by the Simulation Model and is no more to be trusted than an abductive conclusion. These conclusions are, however, usable in several ways: first they do, if in adequate quantity, provide a demonstration of robustness; second, they may be used to abduce the fuzzy boundaries of the issue domain without running very expensive experiments; and finally, they may be used to examine

new hypotheses before subjecting those hypotheses to actual, and very expensive, experimentation.

2.1.6 Validation by Correspondence with Systems Analysis Practice

The computer entered our lives in the late 1940s and became a business tool in the 1950s. Early software applications were simply automations of existing business applications. These programs required very little analysis, just programming. In the 1960s and 1970s software became ubiquitous throughout the business world and with it came the realization that something had to be done to control its cost and quality. The answer that business found was software engineering, a professionalization of the trade of programming. The basis of software engineering was not programming, but systems analysis.

In the late 1970s a number of techniques were developed for software engineering. One of these was Structured Analysis and System Specification (SA), popularized by Tom DeMarco[34]. The technique was designed for and applied to information systems development, but as we will see, that application is a very good analog to any research project managed as a Research Web.

SA starts with a study of the current environment. That study results in the production of a model called the current physical model. In our vocabulary, the current physical model is exactly equivalent to that part of the descriptive model that deals with the physical objects of the system. The next process in SA produces another model called the current logical model. That model is exactly equivalent to the part of the descriptive model that deals with the processes that are observed in the system. As we will see later, the equivalents of the current physical and current logical models are built in the first stage of the research project. In realist terms, we call these models the descriptive model.

The next SA model is the new logical model. As DeMarco puts it, this is the model that represents the system, as it *should* operate as opposed to how it *does* operate. Clearly that is a model of an artifact to be constructed; not, as we might hope, a model explaining *why* the system operates as it does. Yet, in this step of the SA system modeling of the new logical system there are parallels to the realist explanatory model. The new logical model applies business rules to the system -- rules that may not have been well implemented in the observed system. In the realist explanatory model, we might discover *causes* that make it necessary to go back and do some more observation of the issue domain for some more subtle objects or processes that cause the system to act as it does.

When the new logical system is implemented we have a new physical system. This new physical system is exactly equivalent to the simulation model I've proposed. The simulation model runs scenarios with data based on the descriptive model. The activities that are simulated are drawn from the explanatory model and results in an outcome that should be recognizable as some behavior that the real system would or has produced. The simulation model in the realist system would be the equivalent of the SA information system in operation with human activities simulated as well as the information system.

There is little doubt that there has been more effort poured into modeling for software system design and database design than in modeling for all other purposes in the entire history of mankind. Large portions of software engineering modeling practice are excellent analogs of the realist models, validating the realist model by correspondence. In the 24 years since DeMarco's work was published, software engineering tools have been much improved. The object-orientation paradigm and universal penetration of relational databases into information science have resulted in not only a rich literature, but also a great number of computer-aided tools to produce both the graphics and text needed for models.

While accepting that software engineering has shown that modeling is effective and an established methodology, we must take note of the fact that software engineering is focused on the building of information systems. Modeling of reality, an open system, is a much more demanding task than modeling the closed system of a business information system. In our proposed new environment, the Research Web, the descriptive models are more complete than those used to model information systems. Our team members must model a wider range of attributes because they do not have a goal in mind. In an open system we do not know when a relatively obscure attribute may become important.

2.2 Adapting the Existing Research Environment to Collaboration

The existing research environment is the result of several hundred years of evolution punctuated with occasional periods of turmoil when adapting to new technology. It has taken fifty years to adapt to the computer, and now we are in the most chaotic period of adaptation to the Internet revolution. We have, in the past, adopted many revolutionary technologies that remain with us today as integral parts of our research environment. Most adaptation calls for coevolution: the tools of the past are modified to work with new tools that improve the conduct of research. So it is with the environment proposed in this work. The document remains with us, the journal remains, the academic environment remains, older communication technologies remain, and above all, culture and the human psyche remain.

This section examines documents, general systems theory, reward systems, organizational behavior, and communication technology in order to provide knowledge that serves the design of new tools and the organization of research teams to use them. We will discuss the old along with the new and benefits as well as pathologies.

2.2.1 Research as a Literary Enterprise

> Speech is the representation of the mind, and writing is the representation of speech.
> --- Aristotle, *De interpertione I*

It is abundantly clear that science cannot progress without the recorded artifacts of research. Scientific research is framed in documents[35], and scientific progress is made through critical dialog about those documents[36]. While a lecturer can influence the lives of students and change the direction of colleagues; unless the lecture materials are recorded, they will have a very short life. Early scholars wrote books, an expensive medium that did not offer an outlet for short articles. It is probably the scientific journal that is responsible for the industrial revolution and the flowering of science that followed. In the latter seventeenth century the scientific journal began an expansion that is not yet spent[37] [38].

Publication of research findings in scientific journals allows alignment of an author's work with previously published works supporting parts of the author's hypothesis. In a scientific paper, propositions other than the hypothesis being investigated must be supported by reference to supporting documents[39]. The references in a paper recognize the contributions of earlier scientists. Citation analysis[40] is the quantitative study of the scientific contribution of an author as measured by published works and by the appropriation of the author's work by peers. Recognition by peers is the leading reward in science.

The scientific paper is clearly the dominant expression of scientific findings. The current Internet revolution will not change that fact: the genre is embedded in the practice of science. The Internet revolution will, however, change the form of the scientific paper. While the physical representation of the research paper is currently inseparable from the printed page, that is rapidly changing. Many prestigious journals are now publishing

facsimiles of their printed articles on the WWW. Physics is driven by electronic preprints, with journals serving more as archives than as the primary mode of article communication[41]. Usage of electronic journals shows a growing acceptance, but there are reports of a (probably transient) sharp age-based differentiation in usage with 56% of academics under 40 using e-journals, as opposed to only 14% of those over 40[42].

The question of submission of scientific papers to e-journals is quite a different question[43]. Disciplinary and departmental norms frequently deprecate e-journal publication, for good reasons. Several general problems with electronic documents served from the WWW contribute by association to the deprecation of e-journals[44 45]: web-based documents are so easy to copy that originality and authorship is uncertain; ephemerality of the document is unavoidable without long-term institutional support; and revision control methods to prevent multiple versions are not widely enforced. The debate about the relative merits and problems of paper and electronic publication rages on, but the advantage seems to be with electronic publication. Odlyzko[46] suggests that on economic arguments alone, e-journals will become dominant. From the reader's standpoint, cost, accessibility, and utility favor e-publication; while from the institutional standpoint, archival standards, the weight of tradition, and the multi-billion dollar influence of the scientific publication industry favor paper. Lawrence[47] reports that every academic year from 1989-1990 to 1999-2000 showed a substantial increase in the percentage of research articles available online. While Lawrence studied articles in computer science, it seems unlikely that similar increases are not present in the social sciences and humanities.

This dissertation proposes a form of the scientific paper that is more useful than the current printed genre or its electronic facsimile. While no less literary than the printed research paper, the HyperDocument (see §3.4.4.1) remediates many of the shortcomings of the current representations of the research paper. The HyperDocument allows the

reader to quickly access references and associated documents while avoiding some shortcomings of paper-based documents: expense, difficult access, linear organization, and its inability to display animation and sound. Even the argument of portability fades because a HyperDocument printed from the WWW on a color printer contains all the features of a printed document. The HyperDocument is annotatable online, permitting reader criticism of the scientific paper.

Within the RW, there are many other forms of documentation in use. In the detailed work of research, contributions such as critical commentary, definitions of terms, experimental data, opinions, and position papers all serve to advance the argumentation surrounding the research. The RW serves as a repository for all this knowledge.

2.2.2 Research as a Dynamic and Open System

The issue domain of the Research Web provides the focus or organizing principle for the RW, usually expressed as a system model. System models in the social sciences and life sciences are almost always models of open systems[48]. Closed systems exist mainly in controlled experiments, those largely in the physical sciences. The RW must, then, have the flexibility to respond to the unexpected. Even more importantly, the RW must be able to assimilate the unexpected fact or event into its models.

The scientific research most suitable for expression in a RW is long-term large-scale research. This class of research is characterized by frequent changes and additions as research continues. The frequency of revision is directly proportional to the size of the research team and inversely proportional to the amount of current knowledge about the issue domain. Interdisciplinary research is also subject to changes due to discoveries and changes in the dominant paradigm in each discipline. However we choose to look at our knowledge, it is likely to be full of numerous feedback loops[49].

The RW is designed to facilitate the revision of content. The RW Essays can have information appended to them by annotation in DocReview. New literature can be posted to the Annotated HyperBibliography quickly by the facilitator. While preparing a new edition of an RW Essay begins with authoring that is equally arduous in all media, the facilitator has software that makes the re-presentation of the web page quite easy.

With its emphasis on modeling, the RW makes modeling of open systems an interactive and participative design activity. If any member recognizes an omission or error in any model, there is an immediate and obvious way to annotate the model so the changes can be incorporated. Modifying the simulation model programming is a technical task to be sure, but those modifications must be preceded by changes in the graphic/textual descriptive and explanatory models. It is very important to document the presence of influences from outside the current boundary of the modeled open system[50].

2.2.3 Reward Systems in Research

> *Yes, Virginia, scientists do love recognition, but only since Pythagoras.*
>
> --- Leon Lederman[51]

Rewards systems are social constructs. The major influences have been the practitioners of science, followed closely by politicians, ethicists, philosophers and business people. The rewards systems have legacies rooted in the past, and like any legacy has a dominant paradigm and an elite, both of which serve to give the reward systems great inertia. There are two ways to look at reward systems: normatively and realistically. Both approaches uncover desiderata and practical fact. The normative approach gives us the public face of how scientists should be rewarded, and the realistic approach yields a somewhat more sanguine story.

According to the normative approach, the reward system should serve to encourage successful adherence to the four factors in the ethos of science expressed by Merton in 1942[52]: universalism, organized skepticism, communism and disinterestedness. In order to be published a scientific work must adhere to the ethos of science while introducing new knowledge. Only through publication can a scientist gain peer recognition. Merton, in 1957[53], establishes recognition as the jewel in the scientist's crown, with recognition of priority as its brightest facet. Reward systems are central to the study of the sociology of science.

A realistic approach examines personal motivation as the builder of reward systems. Cohen[54] says, "In science, as in so many other professions, the coin of the realm is not collaborative generosity but credit – credit for individuals." The psychology of reward systems is at odds with Merton's idealistic ethos of science. The deepest motivation to scientists may be the desire for power. Power is achieved through recognition[55]. And recognition is achieved by publication of research, but enabled through the exclusive professional structure of science: degrees, awards, memberships and position.

> *Thus recognition is not merely a passive phenomenon. During both revolutionary and normal periods of scientific growth, it is a rite of passage which confers the right to recognize others: it represents a source of power in the scientific community.*
>
> --- Beaver and Rosen[56]

Recognition is based on intellectual achievements, mostly published research that is cited by others, leading to acknowledgement of the origination of ideas. The most effective publications are in peer-reviewed journals of high reputation. The citations must be from established scientists also writing in quality journals. The ultimate reward is to have principles and ideas attributed to oneself permanently[57].

No discussion of reward systems within the collaborative environment can be meaningful without referring to the dominant rewards systems operating today: professional recognition; academic tenure and promotion; and industrial career advancement. The outstanding characteristic of all three systems is that authorship of research papers dominates all evaluation factors. A scientist's reputation rests on credibility which waxes and wanes with the scientist's publications. A "cycle of credibility" strengthens with the volume of quality publications. This body of produced knowledge is the driving variable in obtaining funding for more research and equipment to produce even more publications[58]

While the professional reward system evolved by scientists operates at the highest institutional levels, much of a scholar's career is engaged in satisfying the requirements of the academic tenure and promotion systems. Academia has two or more masters: educating the populace and serving science. There is a great tension in serving masters with conflicting goals. The usual factors for evaluating performance are: teaching, research and service. All take time, but in practice teaching and research compete for the scholars time, while service is largely incidental. Our interest is primarily with research, but the tools and concepts presented here can also serve in teaching and service. Our research will focus on the impacts of rewards systems on research activities.

In defining the nature of scholarly faculty work, Rice[59] defines four forms of scholarship: discovery, integration, practice, and teaching. The first three components are types of research: basic, synthetic and applied. Basic research is new original research, usually intradisciplinary; synthetic research is integrative and interdisciplinary; and applied research is the application of research to the solution of problems. Each of these types of research is treated differently in the reward systems. In academia, basic research is favored, while synthetic research is acceptable though marginal, and applied research is disdained. Boyer[60] points to several forms of scholarship that are seldom recognized,

including writing computer programs and writing for the public press and even popular television. Tool building is simply not rewarded at all[61]. In the humanities, computing is sometime regarded with suspicion and scholars so involved may find themselves marginalized in their departments[62]. Since all three types of research can benefit from the application of asynchronous collaboration in Research Webs, we need to find how the contributions to the RW can be rewarded.

2.2.3.1 Rewards in Practice

The ultimate incentive for participation is reward. Without rewards, participation in a Research Web (RW) will likely wither after the first flush of enthusiasm is spent. Each person on the research team is encouraged by rewards coming from internal satisfaction, peer approval or institutional rewards. Research groups can receive encouragement from peers or from their sponsors or other interested institutions. Rewards reinforce participation while the barriers operate to discourage participation. The response by individuals or groups to the pros and cons is complex and idiosyncratic.

Since participation is central to collaboration and the concept of the Research Web, we must develop new means to reward the team members for contributions other than research papers. We must also attempt to ameliorate the more corrosive effects of current practice. Reward driven behavior as well as the products of such behavior can be used to measure participation in a Research Web.

Contribution of content by authorship is without question the most important single category of participation. The rewards attending authorship are highly developed, though not without problems. Existing rewards for authorship are deficient in dealing with new forms of publication. In the "pecking order" of professional literature, the single author peer-reviewed research paper in prestigious journals is premier. Publications in peer-reviewed electronic journals are often heavily discounted due primarily to the newness of

electronic journals[63]. The most prestigious journals in many fields frequently disregard critical or controversial work and emphasize "institutional" research that extends the current dominant paradigm in the field[64].

Unpublished position papers, working papers, essays and conjectural works that are published on the WWW as part of a RW or personal home page can be important contributions even though they might not be considered for publication by the journals due to length, style, content or prior publication strictures. Writing software for research purposes, even if useful and freely shared with colleagues, is seldom rewarded. Invocation of a scholar's name on the WWW in the RW's public partition or within the RW's working area may become a reward. Study of invocations on the Web may eventually allow alteration of existing methods of performance evaluation by recalling the scholar's name from unusual genres of communication such as commentary, acknowledgments, reports of professional activity, e-mail messages and meeting minutes[65].

Criticism is extremely important; it is the means by which content is progressively refined from a draft to a canonical document. Criticism of a RW essay can eliminate "holes", pose new hypotheses, and contribute to quality. Criticism of references in an annotated bibliography can point out errors and where those errors are corrected. Criticism of the annotated glossary can sharpen meaning of a term or introduce new nuances of meanings. A three-sentence note can change the direction of a paper. Criticism must be rewarded. The RW provides a permanent record of annotations that establishes a record of critical contribution.

Leadership provides the necessary social glue to coordinate the efforts of the research team, to deal with the administrative details, and to keep the work going forward by encouraging participation by the members. While the scientific leadership is generally vested in the Principal Investigator (PI), in practice the PI often shares many of the

functions of leadership. Foote[66] suggests faculty who build or maintain collaborations do not often receive much credit for their efforts. The members who shoulder the day-to-day burdens of leadership should be rewarded.

An analysis of reward systems can proceed from two directions: first is an analysis from the viewpoint of the recipient of the rewards—individual, research team, or cooperating agencies (psychological, sociological, political); second is an analysis of the rewards that may be granted from distinct sources.

2.2.3.1.1 Received Rewards

Individuals are motivated by internal satisfaction, recognition by peers and by the team at large and by organizations cooperating with the research. Internal satisfaction stems from numerous sources: pleasure at the reduction of cognitive dissonance, closure of tasks, learning, etc. Individuals are stimulated by their peers, especially those on the research team; by the team itself, expressed as pride in the team's accomplishment; and to a certain extent by approval of cooperating agencies, often expressed in letters of commendation and in press releases. In the RW, as in other academic groups, social academic rewards are received by two main mechanisms: potential enhancement of reputation, and by reciprocity expressed as permission to use the information contained in the group's knowledgebase[67].

Membership in an active community of scientists is itself a reward. Senior scientists expect this reward, but it is carefully doled out to junior members of the team, post docs and graduate students. This reward is the principal mechanism for socialization of new scientists. Lave and Wenger have named this mechanism *legitimate peripheral participation* and have thoroughly explored the issue[68].

The research team receives rewards of recognition from professional peers in the form of mentions in the press, citation of reports published by the team members, and by communications directed to the team; and from cooperating agencies in the form of additional support in the form of encouragement, additional equipment or support personnel and ultimately the reward of financial grants. Cooperating institutions receive rewards in the form of prestige due to their sponsorship of successful research, recognition of public service, recognition of fulfilling a mandate.

2.2.3.1.2 Granted Rewards

Major rewards offered to individuals by the research team are: inclusion of their essays in the RW, inclusion as author on papers submitted to journals, public acknowledgment of contributions[69], financial support from grant funds, invitation to join the research team, and employment as a staff member. Another class of rewards is ceremonial support that lends official recognition for service and the passing of milestones[70].

The importance of ceremony in team dynamics is often minimized. This minimization may be caused by intellectual hubris. Ceremony does play a large role in our lives and to neglect it is to ask for suboptimal performance. Ceremony is easily overdone, especially when it is obviously artificial. Ceremonies should be sincere occasions that show realistic expectations of participation and the desired effect. Since conventional ceremony always requires physical presence, the distributed research team ceremony faces a special problem: celebration in isolation.

Ceremony in our special environment suffers greatly from the lack of media richness. Lack of personal presence prevents the use of tone of voice, gesture, touching and other forms of communication. We have great difficulty in reinforcing the message of the ceremony by engaging the members emotionally. On the other hand we do have some advantages: the ability to very carefully express one's self verbally, the ability to reflect on the message rather than react to it, the ability of the receiver to recapitulate each

communication, and the suppression of intemperate reaction. If we recognize our weaknesses and capitalize on our strengths, we can successfully incorporate some ceremony into our Research Webs.

Rites of passage, especially those encountered in professional socialization such as graduate school, may be made less traumatic by collaboration with peers[71]. Daniels[72] has teased elements of rites of passage from the ordinary group meeting: Birth (introduction of a new person or idea); Maturation (recognition that a product or project has reached a new stage of usefulness); Marriage (restructuring in the organization); Leadership (promotions, appointments or assignments); Thanksgiving (recognition of awards or acquisition of new resources); Discipline (identification of poor performance and determination of penalties); and Sickness and Death (the identification of problems and allocation of resources to remediation, project, product, or performer termination). While Daniels' insights are useful, Trice and Beyer present a more academic analysis (see Table I below). They have developed a typology of rites that is quite useful for design purposes.

Table I

A Typology of Rites by Their Manifest, Expressive Social Consequences

Types of Rites	Example of rites in the Research Web	Manifest, Expressive Social Consequences	Examples of Possible Latent, Expressive Consequences
Rites of passage	Welcoming a new member to the team. Announcement of retirement of a team member. Rotation of the scientific leadership.	Facilitate transition of persons into social roles and statuses that are new for them.	Minimize changes in ways people carry out social roles. Reestablish equilibrium in ongoing social relations.
Rites of degradation	Dismissing a member and changing group passwords. Removing an author from a writing team.	Dissolve a relationship.	Publicly acknowledge a problem. Defend group boundaries and membership.
Rites of enhancement	Announcement of awards or prizes. Passing on notes of appreciation.	Enhance social identities and their power.	Provide public recognition of individuals for their accomplishments; motivate others to similar efforts. Enable organization to take some credit for individual accomplishments.
Rites of renewal	Announcement of a new essay topic. Announcing an author and team for a paper.	Refurbish social structures and improve their functioning.	Reassure members that something is being done about problems. Legitimate and reinforce existing systems of power and authority.
Rites of conflict reduction	Discussion of conflicts by means of DocReview commentary.	Reduce conflicts and aggression	Compartmentalize conflict and its disruptive effects. Reestablish equilibrium in disturbed social relations.
Rites of integration	Face-to-face meetings at professional conferences. Teleconferences. Conference calls.	Encourage and revive common feelings that bind members together and commit them to a social system.	Permit venting of emotion and temporary loosening of various norms. Reassert and reaffirm, by contrast, moral rightness of usual norms.

(adopted from Table 1, Trice and Beyer 1984)

Except for rites of integration, there is an asynchronous solution for any type of rite. As befits the medium these solutions are all documents. These documents are likely to be

issued by the scientific leader, and should anyone else care to issue them, they should coordinate the release with the scientific leader. Announcement of awards with potential professional value to the recipient should be forwarded to organizations, such as the recipient's academic department, that might find them important.

There are likely to be other types of rites emerge from the genre of the Research Web. For instance, one of the likely events in a mature RW is to develop a spin-off RW to investigate a new idea. Perhaps the RW will be incorporated into a RW of larger scope. Leadership of new RWs is likely to be awarded to outstanding performers, either substantive or collaborative.

Cooperating institutions can grant rewards far more valuable than the offerings of the RW team. Academic departments, colleges and universities may grant tenure, make promotions, award cash grants, change the scholar's duties, or many other coveted rewards. Granting institutions can offer endowed chairs, financial grants or fellowships. Professional organizations or private foundations may grant prestigious prizes. It is important to realize that none of these rewards are likely to materialize without a great deal of persistent effort on the part of the RW leadership.

2.2.3.2 Determining and Distributing Rewards

> *Designing the incentive to collaborate is just as important as designing the technology for collaboration.* --- Michael Schrage[73]

Collaboration is quantifiable in a RW. Most contributions are documented as authored essays, comments and e-mail to the team. Gone are the days when the ability to work in a team is measured by guessing at popularity around the water cooler. The results of evaluation of participation in collaboration may be surprising as it becomes clear how

each member behaves: who takes the time to review a document: who comes up with new ideas; who volunteers to do humble chores; and who offers little to others.

Rewards for collaboration are very poorly developed. At a psychological level, most people feel self-satisfaction with their contributions. The contributions also need to be appropriately rewarded at the sociological level, or at the professional level[74]. In the RW, rewards may be thanks, an acknowledgment in a published paper, or authorship, or may be the first author. Thanks and other expressions of appreciation should be offered consistently as a matter of human courtesy. At a more formal level, the authors of papers should acknowledge everyone who contributes commentary or minor bits of substance to the paper. These contributions are characterized as subauthorship collaboration[75]. The papers produced by the RW team will be part of the RW as essays. There is plenty of room to acknowledge everyone who contributed, perhaps in order of value to the authors. While journal editors may request that acknowledgments be abridged or eliminated, the team controls the RW essay (which may be made public). The acknowledgment is a neglected genre that may, in the coming full-text world, lead to better recognition of collaborators.

As a standard procedure, an advanced draft of any paper that is published should be recast as a RW essay and placed in the public partition of the Web Site. This will provide an annotatable version of the paper. As time has passed, publishers have become more willing to grant permission to allow the placement of an annotatable draft of the paper on the WWW. This will increase the exposure of the authors and those acknowledged.

Authorship is the most visible prize and should be (and almost always is) very carefully determined in advance of the first draft release for comments by DocReview. Additions to the author list, the order of the author list and perhaps the deletion of an author can be

made as the project moves forward. Occasionally a valued commentator may be invited to join the authors. Endersby[76] suggests that all individuals making an active contribution share authorship, while those whose labor could easily be replaced by others, be acknowledged. The APA (American Psychological Association) recommends that only those contributing to the content be awarded authorship, thus eliminating complimentary authorship to heads of departments and laboratories.

The RW team is frequently very independent of academic departments, especially if the team is interdisciplinary. Since the team members are not usually subject to oversight by their academic supervisors, the scientific coordinator should appropriately commend each academic on the team to his or her academic chair. If this is not done, collaboration other than authored papers will likely not be rewarded in the academic department. The prevalent failure of academic administrators to appreciate collaborative behavior is cited in a special study sponsored by the Association of American Geographers:

> *Finally, we recommend that academic administrators and faculty acknowledge that reward mechanisms are now based almost entirely on individualistic conceptions of faculty roles. Without exception in the committee's experience, rewards accrue to individuals evaluated in isolation. That view of faculty roles may prevail for some time. We opine, however, that the kinds of collaborative and team efforts that have proved productive in other industries eventually will prove useful in geography, probably in the form of instructional teams using several complementary methods of instruction.* ***Research in geography also may move beyond the artisan or craft scale that currently prevails, to projects that are addressed collaboratively by organized groups****. [emphasis provided]*
>
> **--- Toward a Reconsideration of Faculty Roles and Rewards in Geography**[77]

2.2.3.3 Penalties

There is a variety of counterproductive behavior that must be prevented. Besides abiding by the normal mores of professional behavior, all participants must avoid or at least reduce certain behaviors peculiar to working in a group. Since most of the work will be done over the Internet, all members should be aware of the rules of "netiquette." Much

more serious is the problem of non-participation or reluctance to participate publicly. Many individuals, especially those in the social sciences and humanities, have been socialized in a world that rewards individual effort. Some people may be shy or embarrassed with their language skill; others may feel that their stature is too elevated to bother with details. Non-participation is a problem that can be solved by three paths: example, training, and leadership. A well-functioning collaboration is an example to newcomers and to team members who are slow to join the party. The facilitator can tutor team members who need help in learning the tools of the environment. Leadership is just as important in a distributed on-line community as it is in person. If a PI lacks leadership skills, the job of scientific coordinator can be passed to someone who possesses the needed skills.

Collaborators who fail to participate can be removed from the team. This penalty may be necessary in order to prevent demoralization of the team. Resentment is a natural reaction toward free riders. David Coleman has established online communities of managers roughly equivalent to Research Webs[78]. These groups have an annual subscription fee for members, and if a member is not contributing, Coleman suggests that these "checkbook members" be asked to resign.

Members who are on authorship teams may be removed if they fail to fulfill their responsibilities. Since authorship is the principal reward in science, such a strong penalty must be used cautiously, with the concurrence of not only the lead author, but the scientific coordinator as well. Another, more effective, class of penalties is admonishment. Admonishments should be considered whenever one of Pröpper's rules of engagement (see Table II in §2.2.4.6.2) are violated. If admonishments are not effective, then removal seems to be the only alternative. Antisocial or disruptive behavior is unlikely unless the RW is open to the public. The team must be able to marginalize or exclude such people[79].

The penalties suggested above are very severe and are unlikely to be applied in an academic environment. Recent research in game theory in economics suggests another control over the free-riding problem. Fehr and Gächter[80] show that if team members are offered the opportunity to punish free riders, they do so, even when it comes at some cost to them. Assume management was to provide a bonus pool to be publicly distributed among the team members based on their peers rating of their collaborative behavior. If the team members were allowed to distribute an allowance to other team members as they chose, then the free riders would be publicly identified by their poor bonus, and of course the prime collaborators would also be identified by a generous bonus.

2.2.4 Participation and Collaboration

Collaboration is, of course, impossible without participation. Participation is developed only through hard work on the part of the team members, and especially the members placed in leadership positions. Participation is encouraged by several mechanisms, among them professional ethos, pay, emotional commitment, curiosity, altruism, recognition and avoidance of the negative consequences of non-participation.

2.2.4.1 Participation as a Covering Term for Interaction

> **Participation**: *n*. 2b. A taking part, association, or sharing (with others) ***in*** some action or matter [emphasis provided]. -- Oxford English Dictionary.

Obviously nothing happens in a group effort unless people interact. The activities are the dimensions of participation. The word *in* links the activity with the object of the action. The object of the activity may be decision making, learning or scholarly research, or any other activity that humans or their surrogates engage in. The principal terms used to describe the *dimensions* of participation are: collaboration, communication, coordination, cooperation and contribution.

Even in the scientific literature, the terms collaboration and cooperation are used loosely. Sanderson notes the close relationship between collaborative and cooperative research[81]. Both forms of participation are working together to a common end, but collaboration calls for a much more dedicated relationship between the participants marked by the creation of new knowledge. Cooperation is a less intense relationship marked by the sharing of resources. The boundary between collaboration and cooperation is seamless. I believe another distinction may be made, by denying that organizations can collaborate, since collaboration is a synthetic intellectual action joined only by individuals. Representatives of organizations may collaborate, but the organizations can only cooperate.

As an illustration of the confusion between the terms collaboration, and cooperation, the "Nairobi partnership" is a pastiche of programs that has grown over the years as a vehicle for studying and attacking the problem of AIDS in Kenya. Some of the institutions involved are universities at Nairobi, Winnipeg, Toronto, Antwerp, Ghent, Oxford and Seattle. Funding comes from various sources and totals over $4.8 millions per year. In the over twenty years of its existence the partnership has produced nearly 300 papers. In a recent Science article[82], the partnership is referred to as a collaboration, yet the Kenyans involved complained at a retreat held in 1998 that key collaboration practices were underdeveloped. Specifically cited was a need to incorporate Kenyans when they draft research proposals, manuscripts, and conference presentations. Also, University of Nairobi staff that currently does not work with the project should have more opportunities to participate.

Of additional interest to this discussion are some parallel relationships: that of objects or symbols of communication -- data, information and knowledge; and terms describing representations -- analogy, metaphor and model[83]. Of particular interest is a modern model of the participating social group: the model of distributed artificial intelligence (DAI)[84]. Artificial intelligence (AI) is an interesting model in that it is based on the reality of individual human thought, yet treats that thought as a network of separated

activities within the mind. DAI adds to the model the concept of interactions between intelligent entities (human or machine).

Why introduce a model using artificial intelligence to the discussion of a very venerable and intelligent human activity? First, DAI is a simplification of the bewildering complex of human collaboration; and second, it allows us to include, as actors, non-human agencies such as organizations, cultures, and yes, even computers. By identification and inclusion of these non-human agents we can better understand the activities of the humans participating in the collaboration.

2.2.4.2 Communication as Participation

> **Communicate:**
> *v. trans.* 2. To impart (knowledge, information and the like); to impart or convey the knowledge of, inform a person of, tell. -- Oxford English Dictionary.

Communication is the process of diffusion of objects in a network of active or passive nodes. The objects that diffuse over the network may be pieces of raw data, information or knowledge. The nodes may be active, as humans are; or passive, as are machines, just responding to requests for information, or accepting information for storage. A library may be considered a passive node that may for instance provide a copy of a paper for a researcher.

Communication has dimensions of bandwidth, or capacity, and path, or connectivity. Bandwidth may be further described as the effectiveness of communication; that would in turn be divided into effectiveness of production, transmission and reception. Bandwidth is a measure of both quality of transmission and volume of information contained in the transmission. Communication is seldom even close to perfect. Communication degrades as thoughts are collected, symbols assigned to the thoughts,

preparation of the symbols for transmission, noise in the transmission, errors in reception of the transmission and finally errors in interpretation of the symbols transmitted[85].

Path is described in terms of impedance of the linkage between two linked nodes. Direct connection is the most effective, and the effectiveness of indirect connections is proportional to the number of intervening nodes and the transformations that may occur in those nodes. Extremely indirect connections are very unreliable and ineffective. Complete lack of connection is possible. Links are directed, that is they are either one way or bi-directional. The impedance in a bi-directional link is not likely to be equal in both directions due to differences in productive or receptive capabilities.

Participation by communication need not be interactive. People often participate passively in discussions, simply reading the interactions of others. This sort of participation is a mode favored by those interested in the topic, but not prepared to engage in the discussion. The reasons for failure to engage may be unwillingness to invest the time required, insecurity (evaluation anxiety), or a conscious recognition that their contributions may not be up to the standards of the group. This behavior is commonly termed *lurking*, and is often wrongly equated with non-participative behavior (free-riding, social loafing, self-censorship, or surrender to authority). Lurking is quite often used to learn about the community, its language, problems, and methods of dialog[86]. Toleration of lurking by the active and elite members of a group is in fact the most democratic example of legitimate peripheral participation[87]. Lurkers are given the privilege of listening to the dialog in hopes that they may eventually join the group as a qualified participant.

Within a Research web, the social ties of the members are strong to weak, but certainly are not non-existent. As time passes, and as the occasional face-to-face meetings occur, the team's social ties will become stronger. Gächer and Fehr note that theory suggests that desire for social approval and avoidance of social disapproval are the mechanism for

maintaining custom[88]. They note the weak empirical support for the suggestion that custom encourages participation; but then show in experiments[89] that while free-riding is only slightly and insignificantly affected by social approval, within a group with at least weak ties, social incentives give rise to large and significant reductions in free-riding.

2.2.4.3 Cooperation as Participation

> **Cooperate:**
> *v. intr*. To work together toward a common end or purpose.
>
> -- American Heritage Dictionary

Cooperation is the act of agreement with the objectives of the participation and the commitment to facilitate the process of reaching those objectives. People cooperate by agreeing to participate, and by agreeing to deliver work in a timely fashion. Organizations cooperate by agreeing to facilitate the participation by supplying data, information, resources, or funding to the participants. Cooperation is frequently promised in the spirit of building an enterprise, but less frequently granted in practice without a contract or memorandum of understanding. Cooperation exists on a continuum of participation that is bounded on the high end by becoming collaboration and on the low end by simply not interfering with the enterprise[90].

Cooperation from academic institutions has, for interdisciplinary projects, often been poor. Disciplinary parochialism has been cited as a problem[91]. For collaborations beyond two-author papers, institutional support is necessary, even if it's only favorable notice in tenure reviews[92].

2.2.4.4 Coordination as Participation

Coordinate:

v.t. To work together harmoniously. -- American Heritage Dictionary

Coordination is a management function. Tasks must be assigned to individuals or groups in order to maximize the efficiency of movement of a team toward a goal. These tasks may need to be related in a network of temporal precedence. Recall that the final action in a task must by definition be preceded by all other activity. Every other act, except the beginning act, has both preceding and following acts. Tasks can be organized in linear or branching networks (serial or parallel). There is an excellent model of coordination, PERT/CPM (Program Evaluation and Review Technique / Critical Path Management) that uses time as its basis.

Coordination is managing dependencies between activities.

---- T.W. Malone & K. Crowston[93]

Communicating task descriptions can result in healthy discussion of the team's overall plans. Criticism of tasks may eliminate unnecessary tasks and point out new requirements. The work performed in coordinated tasks may be thought of as efficient use of resources. To the extent that efficiency declines so does the quality of the product. Very large projects do have plenty of routine work that can benefit from coordination by management. Collaboration, on the other hand, is seldom routine so has little need for coordination.

2.2.4.5 Contribution as Participation

> **Contribute:**
> *v. intr.* To give or supply in common with others; give to a common fund or for a common purpose. -- American Heritage Dictionary

Without contribution of content and resources there is no product. Intentions do not produce results; they are the precedents of production. Contributions are tangible offerings from individuals and institutions to the enterprise. There is a difference in the character of the contributions an institution can make as opposed to the contributions of individuals and groups. Generally speaking institutions are enablers, contributing resources such as services (libraries, computing power) and funding. Team members and their colleagues may contribute resources, but they are primarily contributors of knowledge and the intellectual effort required to bring that knowledge to publication.

2.2.4.6 Collaboration as Participation

> **Collaborate:**
> *v. intr*. To work together, especially in a joint intellectual effort.
> -- American Heritage Dictionary

Collaboration is the creation of new shared knowledge[94]. The collaboration takes place within a context. The context, the issue domain of the RW or the focal topic of a RW Essay, is what transforms information into knowledge. Sharing demonstrates a commitment to the common goal as opposed to personal interests. Sharing information and criticism are critical components of collaboration. Criticism is original contribution: the means by which consensus is reached or alternatives created.

Collaboration is a narrower term than cooperation. People who cooperate come to agreement on goals, but often proceed to those goals independently. Collaboration implies a close relationship and mutual responsibility for the products of their work[95].

2.2.4.6.1 Contribution of Content

Content is King, no matter if the site is a Research Web studying marriage among the Lesser Andaman Islanders or is a commercial site selling shoes. This mantra originated in teaching and is now well established among managers of WWW sites. The Research Web site is a collection of conscription devices[96] for all researchers studying the issue domain. If there is little content, or if the content is static, there is little incentive for returning to the site or for participating. There must be a critical mass of content as well as a critical mass of participants since a small number of participants cannot be expected to contribute at a uniformly high rate.

> *A positive correlation between interests and resources is highly favorable for collective action, as it increases the probability of there being a few highly interested and highly resourceful people who are willing and able to provide the good for everyone.* -- Oliver, Marwell, and Teixeira[97]

The success of the RW depends on universal access, universal adoption, and universal use[98]. Each member must possess the hardware and software necessary to make use of the RW. Each member must be willing to learn how to use the RW by browsing the WWW and becoming familiar with the contents and interactive tools of the RW. And finally, every member must use the RW by criticizing and contributing content, and by actively supporting its use.

Content for the Research Web comes through three mechanisms: commission, acquisition, and reader creation[99]. The RW depends on its members to provide content through all of these mechanisms. Authors of new materials contribute essays, position papers, and research reports in response to a commission. All team members can acquire content for the RW by contributing bibliographic references, and providing definitions of

terms of the vocabulary of the dialog. As readers, the team members turn to criticism of essays, glossing definitions, engaging in e-mail dialog, and reviewing the references.

One early contribution expected from each member is a set of contributions, or perhaps a position paper, that summarizes the relationships of the issue domain to his or her discipline. Basically, this contribution makes the tacit knowledge of the member explicit and open to discussion by the other team members. Researchers engaged in interdisciplinary research projects, as the RW is likely to be, are likely to encourage each other to make the implicit explicit[100].

To hell with tacit knowledge. Go for tacit documents instead. --- David Weinberger[101]

It is very important to remove barriers to the contribution of content. Contributors must not be burdened by needing to become facile in the technology of the WWW. A facilitator must be responsible for reformatting content for presentation on the RW. The contributor should be held responsible only for some representation of the content: word processor file, e-mail attachment, or even hardcopy. Software for direct contribution of commentary and annotation must be designed for novice use.

As one becomes known as a scientist by publishing science, so one becomes a collaborator by contributing content to the RW. Scholarly publications are basically any recorded document that can be accessed by other scholars interested in the topic; similarly, contributions are documents available to the members of the research team. Scholarly publication is a process that has a very restricted set of document genres, so the contributor cannot participate in scholarly publication as fully as one can contribute to the RW with its much fuller set of document genres.

Contributions are organized knowledge and information that include all publications and extend into ephemera such as lectures, performances, and conversation. Contributions of very special importance are documents that exist on the far boundary of conventional

publication: criticism. Criticism in the form of reviews is firmly within the bounds of conventional publication, but letters to the editor are on the boundary, and direct criticism of documents such as might be directed to the author, or discussed in workshops or seminars is infrequent in conventional publication. Yet this criticism is the fuel of scientific progress and the hone that puts the edge on our canonical documents. The reward systems have considered authorship of research papers and books almost exclusively, and have not adequately accounted for smaller contributions, such as collaborative behavior and criticism[102].

Contributions may form the basis of valuation of collaborative effectiveness. The primary difficulties in establishing the value of a contribution are the vast range of the individual contribution and the difficulties in establishing a fungible unit of measurement. Rewards are based on the value of one's contribution. Since evaluation has a large subjective and even political component, it is clear that rewards are not always distributed equitably: how many revered teachers are Nobel Laureates? How many penetrating commentaries equal authorship of a paper? How many annotations does it take to earn an acknowledgment? And, how are acknowledgments valued in a tenure defense?

Commentary that expresses support or disagreement is not valueless, for such commentary does influence the behavior of the author and other contributors. So most commentary is of some value, even if it is merely reinforcing the recognition of a team effort. Sadly there are comments of negative worth that occasionally emerge, such as unwarranted attacks or senseless graffiti.

While valuation of contributions may appear to be a hopeless task, such is not the case. A skilled collaboration leader, if aided by evaluation tools can evaluate the collaborative performance by team members much better than an unaided novice leader. The collaborators and other team members can also contribute to the evaluation. Intangible

bases of evaluation can be incorporated along with automated measurement tools, such as word counts, to develop a well-rounded, largely rational evaluation. This evaluation can be forwarded to the member's employers or can form the basis of letters of recommendation or nominations for awards.

2.2.4.6.2 Collaboration in Development of Content

> *The performance of cognitive tasks that exceed individual abilities is always shaped by a social organization of distributed cognition.* --- Edwin Hutchins[103]

In the Research Web environment, collaboration always results in the development of content. Content ranges from Research Web Essays on the high side, through critical commentary, e-mail communication, reports of meetings or phone calls, to the information gathered in the more tedious tasks of literature research. Certainly all content is not equal in importance, nor is all collaboration on a professional level. All members of the research team from principal investigators to data analysts contribute to the success of the RW, but the rewards and credit are dispensed in accordance with the nature of individual contributions.

The modern research environment is, for three principal reasons[104], a collaborative environment. First, the scope and quality of scholarship has advanced to the point where individual labor is insufficient. Second, collaboration provides credibility, especially for students and less well-known researchers. And finally, the times are right for creation of communities of specialized scholars, and for tightly defined interdisciplinary research topics. The means for collaboration are communication and resources, and the ends are knowledge expressed in documents.

A Research Web must, like all medium-to-large scale human enterprises, be managed. Management is responsible for maintaining an environment that fosters collaboration: all

members need to be granted respect and rewards for what they contribute. All members should adhere to some standards of conduct, as participation is social interaction. In the Research Web, that social interaction is dialog, usually asynchronous. The goal of dialogic behavior in the RW is to attain the ideal speech situation proposed by Habermas and modified by Webler to incorporate competence and responsibility.

> *1. Every potential discourse participant must meet minimal societal standards for cognitive and lingual competence.*
> *2. Every discourse participant must have access to the knowledge needed to make validity claims and criticize others.*
> *3. Speakers must verify the results of any attempt to translate expressive claims.*
> *4. Judgments about conflicting validity claims must be made using the most reliable methodological techniques available.* --- Thomas Webler[105]

Pröpper's Model Procedure for Discussion[106] (Table II below) holds dialog on a professional level to an even more stringent standard. Keeping in mind that research may be viewed as an argumentation process[107], we can adopt a set of behavior elements designed to ensure professional dialog that includes the basis for translation of dialog into formal argumentation. The table below provides rules that not only will bring order into the dialog, but will also enable measurement of the quality of the dialog.

Table II

A Model Procedure for Discussion

1. A committed attitude

1.1 One is committed to the objective of the discussion.
1.2 One is committed to the things one has said and implied therewith
1.3 One is committed to the arguments being solid

2. Accountability

2.1 Every participant in a discussion supports his or her statements with the help of arguments, when other participants (may be expected to) demand this, unless he or she gives plausible reasons justifying a refusal.
2.2 When one doubts the arguments relating to the point of view of another participant in the discussion, one may only challenge these if one gives counterarguments.

3. Consistency

The participants in a discussion act and speak in a consistent way.

3.1 The participants in a discussion are not allowed to contradict themselves.
3.2 The participants in a discussion are consequent.

4. Relevancy

4.1 The arguments one gives, and the information accompanying them, must be relevant.
4.2 when making a statement that (apparently) does not refer to the statements and arguments which are the subject of the discussion, one has to state one's reasons for making this statement, if other participants (may be likely to) expect this.

5. Objectivity

The participants in a discussion adopt an objective attitude.

5.1 One is not allowed to prevaricate.
5.2 One is not allowed to ascribe to another person points of view that one does not support.
5.3 The points of view held must not be tendentious due to ambiguity.
5.4 The participants in a discussion are not allowed to present their own contribution(s) to the discussion tendentiously, by means of incorrect or incomplete information.

Table II (continued)

5.5 One should not become personal.

6. Openness

The participants in a discussion must see to it that the discussion is open to others and to their contributions.

6.1 It must be possible for everyone (to the same extent) to take part in the discussion.
6.2 The participants in a discussion are allowed to raise any point of view and advance any information they consider relevant for the defense or challenge of a certain point of view.
6.3 One is allowed to challenge any statement brought by another participant to the discussion to justify or refute the expression of an opinion.
6.4 The participants in a discussion are to provide as much information as necessary (for the aim of the discussion at that moment).

--- Igno Pröpper[108]

2.2.4.6.3 Collaboration for Labor-sharing

Labor sharing is necessary in any project that requires more than one person by virtue of its size, breadth of disciplinary scope or scheduling pressure. The lone scientist may very well find many projects beyond his means[109]. Research Webs are always of a size that requires labor sharing. Labor sharing applies to all professional tasks: authoring, researching, computing, designing of experiments, designing of statistical analysis, and above all, controlling the quality or validity of the products. Labor sharing is a form of delegation. The delegation of the work is under the control of a scientist who is managing part of the team's efforts. The delegation determines who does the work; the rewards for doing the work seem to flow from a combination of three mechanisms: professional recognition, pay for work and legitimate peripheral participation.

The use of research assistants is almost universal in large-scale research enterprises. Research assistants may perform library research at the direction of the project's scientists. Assistance in data gathering, coding, or analysis is often assigned to research assistants. Volunteer amateurs may perform field research under the direction of the

investigators: archeology is the most frequently cited example, but paleontology and astronomy now use volunteer helpers.

Another form of labor sharing is the utilization of outside technical specialists. These specialists are not necessarily professionally competent in the issue domain, especially where non-traditional tools are employed in the research. Geographic information systems (GIS) are often employed in research in fields not usually associated with the use of GIS. Maienschein cites another example: in 1895 a cytologist was compiling a text on cell fertilization; at that time most illustration was done with artistic methods, but he called on an early microphotographer to provide illustrations of the early stages of cell fertilization[110].

Load sharing in the authoring of RW essays and research papers usually divides the product into sections based on the knowledge and interests of an author. Experiment design and statistical analysis may be shared among the members of the authoring team, and perhaps with other members of the RW team. Reviewing load is also shared among all members of the authoring team. Paper renderings, faxes, and /or DocReview can be employed depending on the quality of the draft. DocReview is especially useful for a "local peer review" where colleagues of the authors and RW members review the final draft before release for publication in the RW.

2.2.4.6.4 Collaboration for Credibility

In interdisciplinary projects representatives of each discipline will provide credibility for their specialty. When a proposal is submitted for an interdisciplinary project review committees can, without regard for the capabilities of the team, reject any proposal that does not have a "certified" member from each discipline[111]. Maienschein cites several examples where the demonstration of community led to increased credibility in the eyes of funding agencies due to a stated or implied compliance with the communitarianism of the Mertonian ethos of science[112].

2.2.4.6.5 Collaboration for Community

Collaboration serves several social functions[113]. For some, engaging the research process with others is simply more enjoyable than working alone. Pre-existing social contacts are maintained and improved through collaboration. Careers can be advanced by collaboration with leaders in one's field. Some perceive that the quality of research can be improved through joint participation[114]. Poole cites interdisciplinary work as a potential savior of thinly populated specialties[115].

In order to survive, every human system needs to reproduce itself. There is a constant turnover in personnel due to death, retirement, or changes in interests. Academic disciplines are a good example of a self-reproducing system, as they place an emphasis on creation of new professors. In the research community, there is a need to socialize new members of the community, and for established members, a need to reinforce the social bonds and professional ethos through collaborative practice. New members of the community (graduate students, post-doctoral fellows, interdisciplinary members of the RW team) are introduced to the operation of the research by a collaborative process[116]. This process is called legitimate peripheral participation.

Legitimate peripheral participation[117] permits the learner to participate in the work of an expert, under the tutelage of an expert, but without complete responsibility for the outcome of the work. Not only learning takes place, but also to some extent, an emotional bond is formed – to science, to the work, to the mentors, and to the team.

Types of social collaborative links

Peer Similar: professional to professional within the same discipline. These scholars may be drawn to participate in order to avoid isolation. Austin and Baldwin call this collaboration "complementary collaboration[118]." This type of collaboration usually begins socially in the discussion of ideas. As soon as a kernel of a unique idea is formed

the collaboration begins to take the form of a research project leading to authorship[119]. The most typical contribution of peer similar collaboration is in the professional dialog operating within existing projects. Peer similar relationships usually operate on the theoretical plane[120].

Peer Different: professional to professional between disciplines. This is a looser collaborative link, but often leads to the same reward as peer similar, authorship. Collaborations among specialists are termed "supplemental collaborations" by Austin and Baldwin[121]. They tend to be loose confederations with limited interdependency. Often the collaboration leads to a project in only one of the disciplines, the other discipline contributing only enrichment. Acknowledgment rather than authorship is often the reward in this case. Thagard points out that, as in peer similar collaborations, the relationships are more on the theoretical plane than the empirical[122]. Peer different collaboration also requires cross-disciplinary education. The RW contributes to this learning by encouraging each member to contribute an essay about the issue domain describing his or her discipline's relationship to the issue domain; in other words, making each member's tacit knowledge explicit and subject to criticism and query. The criticism offered by any scholar will bear the imprint of his or her discipline, offering another opportunity to see the point of view of that discipline.

Maienschein mentions another collaboration combining the two forms of peer collaboration: the Textbook Project[123]. A Research Web could easily be constituted to produce such an artifact, especially if the issue domain was multidisciplinary. A more general name for such collaboration might be the Encyclopedia approach. Collaboration becomes relatively less important than coordination under a single editor. Maienschien points out that the existence of multiple or interdisciplinary authors lends considerable credibility to the project. The RW environment also contains tools to facilitate the enterprise, especially the Annotated HyperGlossary and DocReview.

Professor-Student: Sharing the workload with a student leads to several beneficial results for both the project and the participants. The student can offer skills that are professional, sub-professional, or technical. The professor in this relationship clearly is bound to train the student while the student must contribute to the scientific progress of the team[124]. The mechanism whereby benefits flow both to the student and to the professor, and to the team as well, is legitimate peripheral participation.

Professional-Worker: Sharing the labor load with an employee leads to more speed and allows the professional staff to work more efficiently[125]. Work that can be done by anyone with similar skills needs not be accorded professional rewards[126]; pay alone is sufficient reward, though of course the worker may very well gain personal satisfaction from helping the team move toward its goals. Such collaboration may be called "subauthorship collaboration[127]" because it is usually rewarded by acknowledgment rather than authorship.

Action Research: When the research team is engaged in the solution of a practical problem or is participating in the improvement of an existing practice, such as land use planning, they are likely to be teamed with practitioners[128]. The feedback and reflection between theory and practice produces an enrichment of both theory and practice[129]. Problems of this sort are often described as "participatory" and usually involve intervention, on the part of the practitioners, in an ongoing process. Certainly in this sort of research work, there is an "extended peer community[130]."

“Standard” action research into a problem is practiced by looping through a cycle of fact finding, action planning, action and evaluation. McKay and Marshall suggest that the standard iteration pattern might be shadowed by a complementary research pattern that would provide the backing for action planning with a rigorous development of research questions, hypothesis development, data collection and analysis of results[131].

In terms of the conceptual framework of this dissertation, action research involves a rather radical methodological shift, but no great change in the substantive domain or the conceptual domain. It seems clear that the substantive domain would be much better served by the participation of those researched. The conceptual domain could also benefit from greater participation, though probably to a lesser extent. The methodological domain would be dominated by the research methodologies of participative action research (PAR). The principal impact of action research on the RW would be on the simulation model, if one exists. Since the purpose of action research is to change behavior, parts of a simulation model would be made obsolete every time action is taken to alter the behavior of those researched, the "reference group."

The prototypical Research Web is designed to serve a more conventional research team than an action research team. An RW designed to serve an action research team will need to include provisions for access of those "researched on," those "researched for," and stakeholders who may be affected as well as the researchers[132]. The RW could have facilities for conducting questionnaires and voting on the WWW. Certainly some more "democratic" facilities such as chat rooms and less structured discussion forums could be valuable.

2.2.4.6.6 Collaboration for Posterity

Collaboration for posterity is an underappreciated act. Frequently the foresight of past researchers enables future researchers to engage in longitudinal studies. The data collected in the past was frequently documented in ephemeral documents such as lab notebooks. With the constantly decreasing cost of memory, storage and maintenance of knowledge and data will soon be able to remain on a RW indefinitely. Death, retirement and loss of interest result in the loss of vast amounts of data, information and knowledge[133]; the RW can preserve most of this material.

Because many observations made in the course of research are far below the "minimum publishable unit" threshold, much data is lost forever. The nature of the Research Web is such that there is always room to save observations even though they become peripheral during the course of research. Like a good lab notebook, The RW can preserve false starts and blind alleys. Since all textual documents in the RW are in HTML, they may all be displayed in the browser. Even more important, they are all searchable as a body and individually. E-mail archives, DocReview commentary, RW Essays, all are visible.

The scholarly press has a bias against research that does not show significant statistical results, even though the research constitutes an affirmation of insignificance. In the past, results that were not "positive" were so unlikely to be published that the authors did not even submit them for publication. Fortunately, funding agencies such as the National Science Foundation now demand that data and reports from "inconclusive" research be made available. In the Research Web, there is a prominent place for such results. These studies can be made available not only for access to results and data, but to annotation from other scientists that might suggest alternative hypotheses or methods.

2.2.4.6.7 Collaboration through Criticism

The important place of criticism in science as a basic epistemological tool is very well established by scholars such as Popper and Polyani and is far beyond argument in this dissertation. For a philosophical discussion of the issue see Miller[134]. Criticism poses problems that stimulate research and open new directions and keeps researchers from becoming complacent. Criticism can support theory by showing the errors in competing explanations.

> *The best way to advance knowledge, it follows, is to foment a constant stream of criticism and response.* --- Marshall Scott Poole[135]

Critical social theory suggests that research include all those affected by the research. This call to collaboration also clearly states that criticism cannot be separated from

reason[136]. Critical social theory also points clearly to the use of action research as a method for effecting change through inquiry. The Research Web concept is fully compatible with these ideas.

The Research Web is designed to promote criticism: of assertions in essays through DocReview; of definitions and glosses in the Annotated HyperGlossary; and of the quality of references in the Annotated HyperBibliography. Responses to the criticism may be made within the critical apparatus of the tools or by publishing new editions of the essays. As a matter of course, new editions should include a hypertext link to a "preface" that summarizes the changes made in the new edition. As criticism is received it is announced to the members automatically by notification services in DocReview and "What's New." Any e-mail dialog on a given topic can be extracted from searchable e-mail and discussion forum archives.

2.2.4.6.8 Mandated Collaboration

Granting agencies and other organizations that sponsor research often have political policies that mandate collaboration between researchers, disciplines, or organizations. During the 1920's the Rockefeller Foundation was instrumental in encouraging a collective, communitarian, attitude in science[137]. More recently the National Science Foundation has very explicitly sponsored collaboration in its grants[138]. NATO's Collaborative Linkage Grants require collaboration between scientist from NATO countries and certain specified countries in Northern Africa, the Middle East and the former Soviet Union[139].

In 1945 the United States created the Atomic Bomb Casualty Commission to study the health effects of the Hiroshima and Nagasaki bombs. The collaboration of Japanese professionals was essential to the project, yet the cooperation of the victims of atomic war was not going to be obtained without considerable political, sociological, and

psychological manipulation. In short, the occupation authorities mandated the Japanese to collaborate[140].

Collaboration in an externally mandated collaboration is influenced by the formal position the scientific leader, PI or convener may have been invested with by the mandating agency. The participation of the members may be insured to some degree by what they might lose if they do not participate[141]. Internal mandates are also imposed by disciplinary practice or professional ethics[142]. Participants are also bound by the ethos of science and funding obligations.

2.2.4.7 Cooperation in Supporting the RW

Cooperation is the act of supporting first the objectives and operation of the RW and secondly, the support of the members of the team. The sources of support are either institutional or personal. Those with the means of production (land, labor and capital) must support the production of content. In our context, office space and a WWW site represent land. Labor includes the members of the research team and all their supporting staff. Labor must be rewarded for their efforts. While payment in intangible rewards is sufficient for those who spend little time on the research, salaries and office space must support those directly and seriously involved. Capital expenditures may be needed to purchase equipment or to hire consultants or special services, such as laboratory procedures. In University-based research these contributions are all delineated in the research proposal. For long-term collaborations, means other than grants must be assembled. Endowments may pay for laboratory operation, or for intellectual involvement. But for long-term, low budget research efforts, the leadership of the research team will depend on intangible rewards to the intellectual contributors and on the good will of cooperating institutions for facilities.

2.2.4.7.1 Institutional Support

Institutional support of the RW can be financial, persuasive through sponsorship or endorsement, or in the provision of resources. Members of the team may be supported with salaries or grants, lab and office equipment, offices, and privileges of use of facilities such as libraries and computing. Institutions benefit from support of successful Research Webs by gaining stature or fulfilling some institutional goals or mandated duties.

The distributed nature of the RW provides some unusual resource sharing opportunities. Software licenses are often granted to institutions without restriction to the number of users. If a team member is an employee of an institution that has a license for a specialized piece of software, the RW can utilize that institution's WWW server and the software license to support the RW. For example, one major feature of the RW is the e-mail discussion group or list server. The list server software supports the discussion group by distributing e-mail, and by maintaining archives of the team's e-mail that may be searched from the WWW. That list server can reside at any institution with a site license. Courtesy accounts may be provided to team members from other institutions in order to give them access to computing power and on-line journal access.

Institutional support for the members of the research team is excellent in most industrial enterprises. University supported grants, on the other hand, usually are awarded to individuals or very small groups for the purposes of generating narrowly focused scholarship that is usually monodisciplinary. Departmental support for scholars collaborating outside their department is often lacking, even to the extent that the time required for outside collaboration is expected to be done "on the scholar's own time[143]", or as a "night job[144]". Faculty rewards for working collaboratively on the WWW are few, even though in the opinion of some collaborative projects are where the future lies[145].

2.2.4.7.2 Personal Support

Personal support of the RW can be in the form of contribution of intellectual capital, criticism, resources, or endorsement. Personal support of team members can take the

form of recommendations, tutoring, advice, assistance, or simple encouragement. Individuals benefit from support through association and acknowledgment, not to mention the considerable value of personal alliances.

Members should familiarize themselves with the resources available through their institutions. Often site licenses are available that can be used for the team's benefit through a member's institution. Members may have personal resources and abilities that can be applied to the team's efforts, for example the ability to compile programs in "exotic" languages, convert database formats, write programs, or create graphics.

There are three sources of personal support for team members: the scientific leader, the facilitator and each other. The scientific leader has a support role to play, one of assisting and encouraging the team members on a personal level. This sort of support is basic leadership and the ability to exercise such leadership is a basic qualification for the scientific leader. The facilitator is usually not a resource for scientific content, but should be able to help team members learn the methods required for effective asynchronous communication. This help should extend beyond teaching the use of the special tools such as DocReview and the Annotated HyperGlossary and HyperBibliography. The facilitator should be able to detect problems that a novice is experiencing and then tactfully assist the novice in learning even the most basic Internet skills. Before embarking on a mission of assistance, the facilitator should discuss the problem with the scientific leader so as to avoid potential "pride" problems. All members need to be aware of any adjustment problems anyone is having with asynchronous methods.

2.2.5 Barriers to Collaboration

> *... the greatest barriers to development of effective collaborations on the World Wide Web may revolve less around advances in the technology itself than around the institutional dynamics of higher education—human and organizational barriers that are more difficult to change.* --- Kenneth E. Foote[146]

2.2.5.1 Ownership

The commerce in ideas and services functions through exchange just as commerce in goods. Intellectual commerce is a great deal more complex than trading in goods. Just what are the objects of intellectual commerce? Who owns these objects? Who can own these objects?

Authorship of research papers and books belongs to those who wrote the work. Copyright generally is held by the author(s), though it may by contract be assigned to an employer or publisher, or to the RW if it is incorporated. Acknowledgment of contributors is the scholar's courtesy[147].

Models are almost always the work of the authors, but if not they should be cited and/or acknowledged. The diffuse 'ownership' of the models represents a serious threat to participation in the modeling process. As the organizing 'glue' of the RW web site, the models are communal property. The models, much more than products such as the Annotated HyperGlossary and HyperBibliography, are the products of the entire team. The RW Essays are the product of authoring teams, and the rewards will go primarily to the authoring team. Authoring teams, especially solo efforts, may be reluctant to contribute their submodels to the communally held models. Scholars that see no personal rewards from modeling may resent the granting of scarce resources to the construction and maintenance of the models. Short term thinking by authoring teams that choose to terminate their research efforts after the research has produced most or all the papers it is

likely to inspire will result in damage to the modeling effort and also to the quality of the research articles by truncating the development of robustness (phase 3 of VNS, see §2.3.3.3).

Unpublished essays, substantive ideas, RW essays and annotations are owned by the author(s) on the byline. Copyright defaults to the author(s) by default. Common courtesy on the part of scholars demands that important ideas be attributed to the person who suggested them. Ideas may be cited by URL if expressed in e-mail, as personal communications, and/or mentioned in an acknowledgment section. In the RW, if annotations are made in DocReview, important critical contributions cannot only be attributed, but can be made available in full text. The RW records provide better provenance (URL) than the "personal communication" citation treatment usually given this material.

In the RW, maps, charts, graphs, photographs, other images not made by the authors can be attributed to the makers. Materials of this nature need not be attributed if done for hire. Due to the hypertextual nature of the RW, a sidebar may be made available that discusses the technical aspects of the image, including metadata, interpretive notes, and warnings. The maker may separately copyright images.

Software can be copyrighted and sometime patented. If done for hire, the sponsors may copyright it; otherwise the copyright belongs to the programmer. Certainly software critical to the research must be cited and/or acknowledged. Algorithms cannot be patented, but they can be acknowledged in the software code. With patents, the inventors of record are those who were responsible for the development. Acknowledgments are not made in the patent documentation. Ownership of patents is usually assigned.

It should be understood by all contributors that, unless otherwise stated, all commentary and email directed to the team's listserver is open to the team. If one does not wish to make any communication public, then normal communication channels are available for private use. In the normal functioning of any social group there are communications that need to be private. Constant resort to private communication within the research team is somewhat pathological. One function of the scientific leader is to lead reluctant members away from private communication of information that should be public. Not only is such communication an attack on the efficacy of the work of the team, but leads to the establishment of cliques within the team. There are significant psychological mechanisms in operation within the research team, such as the impostor syndrome, and evaluation anxiety[148].

Most documents mounted on the RW have a byline. The byline establishes the ownership of copyright of the document. The byline also establishes "ownership" of an idea within the team, if the document is substantive. Ideally, a white paper will establish a topic for a RW essay. The team's commentary of that white paper should give the author some idea of who might be co-authors. As soon as a publication is planned, the author(s) and scientific leader need to establish who are the authors and the order that their names are to appear on the publication. Authors may later wish to drop out, and others may be added.

Unless legally incorporated, the research team has no legal standing. The team members cannot jointly own any copyrights unless all members agree to jointly register copyright on those documents. Copyright automatically defaults to the authors listed on the byline. The models produced by the team are the embodiment of the long-term research into the issue domain as contrasted with the constituent topics that are the subject of essays produced by authoring teams. The models should be copyrighted to the entire team, past and present.

2.2.5.2 Copyright

At the time of writing copyright practice is in great flux, even anarchy. The law is far behind the times as it was largely drafted in the pre-internet age. We are not concerned with the obvious violations such as plagiarism, "mirroring," and unauthorized excerpting, but in the more subtle area of technical illegalities that are almost universally tolerated.

Copyright is extremely important to the Research Web, as it is likely to include copyrighted materials. The Annotated HyperBibliography is in violation of copyright law unless public availability of its abstracts is permitted by each of the abstract copyright holders. The Annotated HyperGlossary is likely to contain verbatim copies of dictionary definitions. Only a completely private RW can use the abstracts or definitions under "fair use" laws. The argument that a RW is closed to all but the research team is insufficient, as fair use is generally applicable to individual scholarship and teaching.

With the revolution now underway in scholarly publishing, copyright customs are being challenged by universities. For decades scholars have surrendered their copyright to journal publishers. Research reports produced by public funding have been ceded to publishers who then have the right to charge the public for copies of the reports that they have paid for with their taxes. Both the government and the Universities are beginning to question the status quo.

Publications arising from the RW cannot be mounted on the RW without permission of the copyright holder. This absurd situation can be easily resolved by obtaining the permission to publish on the RW prior to submission to the target journal. In the author's experience such permission has never been refused, as it is made clear that the document is a "highly augmented hypertext version of the research report that is likely to change frequently."

2.2.5.3 Tenure and Promotion

Academic rewards are, especially for junior faculty, focused on the tenure and promotion (T&P) process. These rewards are structured to serve the academic department and the college, not the scholar, students or state. T&P are to some degree determined by score-keeping formulae. The formulae are strongly entrenched for several reasons: custom, ease of administration, and (in the United States) the presence of a defensive strategy to avoid discrimination lawsuits. With a mixture of publication categories and quality weights, a basic score for scholarly research is computed and then merged with scores for public service (outreach) and teaching to determine the eligibility for tenure or promotion. The prevailing perception of the importance of research generally leads to the devaluation of other goals of the academic institutions[149].

Solo publications are, other things equal, usually rated above multi-author publications regardless of publication quality[150]. Review groups generally accord low value to being third or more down the author list[151]. While single-author publications may arise from the RW, most publications are multi-author efforts. Acknowledgments, almost obligatory in RW works, carry no weight whatsoever in most T&P formulae[152]. Ruhleder reports that in the humanities, though tool-building is a scholarly activity, "developing computer-based tools is not even on the list[153]." But the most corrosive barrier to collaboration erected by academic departments is the stricture to publish within the discipline and in the leading journals of the discipline[154,155]. Interdisciplinary work published outside the discipline's journals is somewhat deprecated in T&P formulae. The pressure on junior faculty to publish quickly and frequently tends to promote shallow work[156]. Due to the limited time available to build a winning tenure case, few junior faculty members have the inclination to risk becoming involved in groundbreaking work[157,158]. This sort of work is, of course, just the sort of work that usually requires interdisciplinary collaboration.

Collaboration itself is not valued by T&P formulations[159]. Building and maintaining collaborations on the WWW, or helping to build the necessary infrastructure are not often rewarded[160]. Bohen and Stiles are of the opinion that participation in collaborative enterprises are not usually factored into the scholar's workload, so such participation must be "night work[161]."

2.2.5.4 Institutional Barriers

Each discipline nuances the language used in its science[162,163]. Terminological problems are one of the easiest disciplinary constraints to overcome[164]. Understanding the meanings assigned by different disciplines can be a long process. The author sat through two hours of meetings with geographers, ecologists, statisticians, worker safety specialists, and toxicologists that discussed a single term—hazard. Interdisciplinary collaboration may also be constrained by disciplinary methodological biases[165166].

Journals, with the exception of the few major general journals, are designed to serve the needs of a single discipline. This can introduce barriers to publication of collaboratory and especially multidisciplinary research[167]. Journals may have a policy restricting multiple authorship, clearly penalizing collaborative work. Multidisciplinary research may be seen to not be "cutting edge" research, and publication of articles outside the discipline's dominant paradigm may threaten to lower the journal's prestige. Journals in sociology and psychology are "hard to crack" for scholars outside those disciplines[168].

In some disciplines, especially those in the humanities, scholars are socialized into a culture that customarily performs solitary research[169]. This isolation promotes secrecy and competition rather than openness and cooperation and collaboration[170]. This isolation and competitive environment pervades the scholar's life from primary school through undergraduate days in the form of competition for grades. The struggle for tenure is usually a solitary extension of graduate research designed to yield enough

publications, preferably single author[171], to fulfill the requirements of the T&P formulae[172]. Isolation is augmented by the new scholar's socialization into the department, not into the college, discipline or the larger society[173].

Disciplinary boundaries are neither eternal nor eternally useful. --- R.L. Kahn[174]

Damrosch uses the cultural myth of the isolated scholar to drive home his points regarding the corrosive effects of academic culture on collaborative activities[175]. He points out the natural tendency of the young to collaborate and the institutionally encouraged collaboration that pervades learning from elementary school through undergraduate education. Suddenly, upon entry into graduate school the student is forced into the isolated scholar role. Collaboration takes a new name—cheating. Should the student earn a professorial appointment upon being awarded the Ph.D., the six or eight years spent earning tenure reinforces the mold. The isolated scholar often develops a tendency toward secrecy rather than urges to collaborate. Accompanying the myth of the isolated scholar is another myth especially prevalent in the humanities and social sciences: that of mentalism. Mentalism proposes that all good ideas and the publications that come from them can only be the product of a single mind[176].

Collaboration has suffered greatly by academic policies that have sometime produced effects that act as barriers to collaboration. Colleges and academic departments have created and maintained rules and policies that serve their interests as administrative units and instruments of implementation of public mandates such as teaching, research and public service. Unfortunately, they are often not questioned due to their canonical nature. Drawing from official documents from their University, Ervin and Fox point out numerous policy statements, both subtle and frank, that discourage collaboration in both theory and practice[177].

Academic culture in the humanities works against student/professor collaboration in three ways[178]: grant funding is often meager, allowing no support for graduate students; research in the humanities does not lend itself easily to division of labor, in sharp contrast to the social and physical sciences; and finally there is a strong tradition of solitary research in the humanities, most scholars did not work with their advisors, so they seldom work with their own students.

2.2.5.5 Scholarly Competition

Since priority in publication is such a powerful prize, there is an understandable tendency to hide certain key pieces of information. Career advancement goals may become a disincentive to information sharing. Individuals sometimes lose sight of the team goals and press for incorporation of their own expertise into the team's research[179]. Brody suggests that scientists need an easy technique for making some components of their work publicly available while making key components available to a restricted group[180]. Such discrimination in information access is now available by depending on preprint distribution networks. In a RW, a public partition can be created for public information release, and restricted information can be secured by utilizing private e-mail or a passworded partition of the RW (the team partition).

2.2.5.6 Funding Mechanisms

Funding by governmental agencies dominates support for research. Unfortunately funding by these agencies tends to support narrowly drawn goals that can easily be achieved within a short time period. Reviewers for these agencies also frequently disallow the added expenses of sustained collaboration[181]. Short term goals and lack of support for collaboration doubly penalize research webs. Turning the RW's long-term large-scale outlook into a positive attribute is a challenge to the creative grant writer.

Committees of peers award grants, and these committees have topical biases. Jared Diamond points out disciplinary bias in grants awarded to researchers in mental disorders[182]. The National Institutes of Health award many grants for biochemical work, but only a few for 'talk therapy' despite the obvious success of counseling, for example. Fortunately for collaborative teams, there is a positive bias toward collaboration. So we see just what one would expect from a committee of peers: tendencies to select research that supports the dominant disciplinary dialog.

2.2.5.7 Human Culture

National character affects collaboration profoundly. The principal support for this proposition is the work of Geert Hofstede that investigated IBM employees throughout the world[183]. In a survey that produced around 117,000 answers, Hofstede found significant differences in cultural behaviors on four theoretical dimensions: power distance; uncertainty avoidance; individualism; and masculinity. Another dimension was added in 1984: Long Term Orientation/Short Term Orientation[184]. This dimension was found to differentiate Asian workers from the "western world", and was for a time referred to by Bond as "Confucian Dynamism".

National character is molded by the culture of that nation. For example, in the United States individualism is instilled into the children, especially males, of the society at an early age and reinforced by the national mythology. Collaboration and cooperation is also damaged by competitive customs represented in speech as Machiavellian clichés such as "Knowledge is power," and its corollaries "Don't volunteer information" and "Don't give anything away." The pervasive influence of Confucianism in Asia has been cited as an obstacle to the practice of science[185]. Confucian training teaches the student to accept the training of his masters, leading to a scientific inertia. Hsü describes an epiphany received when he found himself disbelieving physical evidence that passed before his eyes[186]. The evidence, drill cores, supported sea-floor spreading, the primary evidence supporting the then new theory of plate tectonics; that theory was in conflict

with the teachings of his professors. Within a given pluralistic nation, religious and ethnic communities, and class differences create a mosaic of attitudes affecting collaboration. For instance, in British Columbia, some native North Americans practice "information bartering[187]". Asking a question is likely to be met with evasion unless some information is offered. Harvey, in comparing Geographic Information System installations in the USA and Germany, attributes the Germanic propensity for regulation to Hofstede's observation that Germanic cultures exhibit high risk avoidance[188].

Bantz isolates four empirical difference factors that affect intercultural interaction, all embedded in communicative patterns: Language, Cultural Norms, Status, and Politics[189]. Arguments have been put forward that occupation is a greater determinant of behavior than nationality[190] but Hofstede's data is based on behavior in a corporate culture that is a classic of conformance. I certainly would accept the degree of conformance in the international scholarly community somewhat greater than in the corporate culture of IBM. Heaton points out that occupational culture and national culture are distinct and interrelated[191]. Kurland and Egan suggest that engaging in dialog on the Internet requires a culture of accountability[192]. I do expect to see cultural differences affecting the work of Research Webs very significantly.

Bantz suggested some resolutions to the four empirical difference factors. His tactics for dealing with the language facility differential are designed for synchronous communications situations, but include one very usable in asynchronous communication: restating concepts in native languages[193]. Such restatement could find a natural place in a DocReview of a document. "Off-line" conversations between bilingual members and a member that may have language difficulties would be most helpful. Members should be sensitive to language differentials and should avoid colloquialisms and metaphors that may not convey meaning to non-native speakers. The Cultural Norm of work norms[194] does not have as great an impact on asynchronous proceedings as it does on synchronous environments. Nevertheless, all members should be forthcoming about when they work

and when they rest. Tactics for managing Status Differences[195] are centered about making every member aware of the status of all the members and how that status might affect group communication. Fortunately asynchronous communication reduces the effects of status differential.

There is also the very important question of the differential behavior between synchronous and asynchronous social interaction. The absence of turn-taking conflict in asynchronous interaction certainly must blunt the power of status. It has been shown that e-mail weakens the power of social status[196]. The ability to reflect at length on one's communication and the communication of others certainly reduces the power of the native reader over non-native readers; this in stark contrast to the power of native speakers in face-to-face argument[197].

Conflict resolution presents major problems. Cultures that have a high level of respect for authority may abdicate their positions in the face of high status. Bantz observes that the behavior of members who are open and direct may introduce social difficulties when interacting with members who are oblique and indirect[198]. DeMente points out several ways how Japanese salarymen defer to authority and to the group[199]. In the Arab culture language use is marked by indirectness and elaborateness that may frustrate those from Anglo-European cultures[200].

The cultures of the communicants are important because the cultures determine the weekly cycle of activity and modify that cycle by mandating holidays determined by local religions, regional customs and national histories. The culture of the communicants also modifies the diurnal cycle by setting the length of the workday and by inserting customs such as extended mid-day mealtimes. The physical location of the communicants is important because human beings synchronize their activities to the position of the sun in the local sky. So the physical geography of synchronous communication largely determines the diurnal cycle of the collaborators, and its human geography studies the

cultural modification of the diurnal cycle, the weekly cycle of the business activity and the annual cultural cycle of holidays.

2.3 A Methodology for Framing a Collaborative Research Process

The purpose of this section is to show how the Research Web (RW) can facilitate most aspects of the research process. To do this we must bring some order to the bewildering diversity of research as practiced. The VNS (Validity Network Schema)[201] provides a complete general description of process flow for an idealized research project. It is this ideal research project that we use to demonstrate the utility of the RW.

The VNS, a multi-perspective framework for understanding the inner nature of the research process, is one of the three foundation concepts in the conceptual framework for this dissertation. Below we will examine how the Research Web environment can facilitate the research effort from the VNS perspectives of: *stages*, or the temporal progress of the research; *paths*, or the conduct of research leading to a specific product; and *domains*, the basic epistemic thrust of the researchers.

2.3.1 VNS: An abstraction of an ideal research process

Research in practice is approached from a bewildering array of models. Each discipline has its own models for reports, and its own set of preferred models, a recipe. Research by recipe takes the researcher's mind away from the meaning of the research and replaces it with a set of tasks. VNS, in its ethereal abstraction, leads the research team to results by method rather than by task performance. VNS is well suited to any size of research team since the conduct of research must suit the behavioral nature of the individual team members. Unless the team is composed of "renaissance scholars," less knowledgable persons must join hands with others to explore each of the several approaches to the research. VNS depends on knowledge building from multiple points of view, and the RW organizes that knowledge.

The ideal research team needs three kinds of knowledge specialists: theoreticians, empiricists, and methodologists. Each of the team members is expected to team with members from the other specialties in order to produce new knowledge, or products. The interdisciplinary nature of many research issues suggests that scholars from disciplines outside the primary discipline can contribute to the knowledge building[202 203]. There are a number of functional roles to be fulfilled in the RW. These roles can last for the lifetime of the RW, a facilitator is always needed to reduce the cognitive load on the researchers; or they might be of short duration, as the need for a very specialized statistician for analysis of experimental results. The team members can move between roles, or researchers may serve on the team for the duration of only a well-defined task. While the RW is not a management scheme, it is an organizational shell flexible enough to adapt to changes in leadership, retirement, loss of interest, and other perturbations.

One role is pervasive: the role of collaborator. Every researcher must be, or become, a collaborator. Colleagues from other disciplines may be asked to participate, as a contributor or collaborator, in information gathering, evaluation of products, or document review. The role of author will persist only for the time needed to produce the document or research report. Authors of research essays may be asked to manage a team of contributing authors, and then be responsible for hosting a collaborative dialog leading to a succession of refining editions of the essay. The role of scientific leader (PI) may rotate, or may be shared. The role of facilitator is tangential to the research goals, but must be filled with a person, or persons, who have an interest in the research topic, and above all a willingness to serve the research team. The facilitator should teach members of the team how to use software that they might find useful. The facilitator should also train his or her replacement well in advance of departure. The role of critic is essential to maintain the quality of the RW. Criticism of essays is the engine of refinement. No document, reference or definition is exempt from criticism.

Due to the size of the RW team and the large volume of information and knowledge brought to bear on the research issue, there is a correspondingly large body of organized knowledge produced. The body of knowledge, the RW repository, can be easily used to develop multiple hypotheses, and thus will ultimately produce a stream of research results until the issue domain is well known. The RW is thus seen as an incubator of issue specific research. From this set of facts springs the necessity of long-term research. If the issue domain is dynamic, the RW might usefully persist for more than a lifetime.

The VNS applies throughout the range of RW applications between the small-scale single paper collaborations and the very large scale confederations of research efforts such as The Cochrane Collaboration[204], a very large organization devoted to meta-analyses of evidence-based medical studies, and the NCGIA (National Center for Geographic Information and Analysis), an incubator for diverse research efforts. Small-scale collaborations cannot afford the overhead involved in the RW, and confederations suffer from a low level of interdependency, a quality essential to participation[205]. Note that large-scale research confederations can include a set of independent RWs joined by a Web site that describes the purpose and emergent qualities of the confederation and provides an introduction and index into the Research Webs where the work is done.

2.3.2 The Research Domains

VNS divides all the knowledge and research effort in the issue domain into three domains: conceptual, substantive, and methodological. Each of these three domains attracts researchers with a preferential point of view, either theoretician, empiricist or methodologist, respectively. Of course no scholar can avoid taking part in all three points of view in the work toward an advanced degree. Still the preference remains, part of the personality that frequently becomes a hallmark of the scholar's reputation.

2.3.2.1 The Substantive Domain

The substantive domain contains a body of existing knowledge and new research designed to capture the observed nature of the objects and processes of the issue domain, the focal topic of a research project. It is the reality of the issue domain that exists prior to our research[206] and is the grounding for the theory and methodology, such as EAST2[207], that we might apply to our research. New observations made during the research will also find its way into the substantive domain. The substantive domain cannot simply be a repository of isolated information about the issue domain. It is so large that it needs to be indexed and organized. The organization is the function of the descriptive model. The descriptive model will describe all the objects in the issue domain and all the observed processes.

Empiricists work in the substantive domain. These researchers are primarily interested in discovering and describing the workings of some part of the issue domain. Their work includes the minute detail that others with a more grand view disparage as "stamp collecting," forgetting that grand visions are based on collections of observations. I think of naturalists like Linnaeus and Nikko Tinbergen, and the early astronomers such as Tycho Brahe and Charles Messier who built great catalogs as stereotypical empiricists: the observers, measurers, classifiers and catalogers. Brinberg and McGrath[208] add to that the system specialists and practitioners, perhaps the people that Funtowicz and Ravetz[209] call the “extended peer community,” and those Donald Schön describes as professionals in his work on reflective practice[210]. This large group of people would include engineers, planners, nurses, educators, social workers and many other experts.

2.3.2.2 The Conceptual Domain

The conceptual domain abstracts the objects and processes of the focal topic, as described in the substantive domain. This abstraction characterizes the nature of the observed relationships and attempts to provide a causal explanation of why the issue domain works

as it does. Each of those causal relationships, or mechanisms, becomes a hypothesis. The emergent explanations constitute the development of theory.

Theoreticians work in the conceptual domain. Basic science is practiced in this domain. Geographers such as Christaller and Lösch created central place theory in geography, and Wegener, a meteorologist, created the concept of "continental drift" that eventually became the theory of plate tectonics. The conceptual domain acts as a hypothesis mill, always turning up new research questions to examine by experiment. While scholarship is needed and practiced in all three domains, interest in the conceptual domain characterizes most scholars.

Theory can be appropriated as well as developed. A theory, EAST (Enhanced Adaptive Structuration Theory), was developed in 1996 to explain collaborative use of Geographic Information Systems for complex decision support processes[211]. EAST was then appropriated for a larger study in a more comprehensive study of participative GIS[212], and evolved into EAST2. Whether theory be appropriated or developed, new patterns from the substantive domain need to be incorporated into the conceptual domain. New patterns can lead to confirmation of theory, to extension of theory, or to correction of theory.

2.3.2.3 The Methodological Domain

The methodological domain contains all information regarding the treatment of the data gathered to support the study of a hypothesis. It will contain information regarding the operationalization of measurements, the measurement protocols, and the techniques to be employed to analyze the data to support or reject hypotheses. The nature of the variables, the operationalization chosen, and their relationships will determine the mode of treatment[213]. Very loosely, this can be referred to as the basis of research design[214].

Methodologists such as statisticians and experimentalists dominate the methodological domain. In the social sciences, measurement of a studied variable often requires selection of a surrogate that can be directly measured. These decisions, known as operationalization, comprise the auxiliary model[215], a model of the system that is actually measured. The RW documents the auxiliary model by making hypertextual links between descriptions of the operationalizations and the model of the system as described in the descriptive model and the explanatory models. The RW will contain several documents that describe the design of the experiment, the operationalization of variables and the analysis plan. The team members can place these documents in DocReview for local peer review.

The RW may, in mature sites, contain a simulation model of the issue domain. A simulation model can be started when the descriptive and explanatory models are sufficiently complete. The simulation model can be used for validation of the explanatory model. Since a simulation model is hierarchical, completed and validated submodels of the system may be published or placed on the RW's web site to be executed by anyone. A working simulation model will allow the research team to perform sensitivity analyses on the variables of the model. The sensitivity analysis will determine which variables contribute most to the variance of model output. Sensitivity analysis can also be used to evaluate determination of the relevance of mechanisms of submodels, in other words, can determine if certain elements can be removed from the model. The team can run scenarios that can examine unobserved situations and create plans for experiments to refine the model.

2.3.3 The Stages

VNS suggests that research studies have three stages: stage 1, a prestudy, or proposal phase; stage 2, a central research effort; and stage 3, making the research findings more robust. Stage 1 has well defined milestones to mark its progress: the proposal, receipt of funding, and completion of a research plan. Stage 2 has less conclusive milestones; research papers might be released during and after the central study. Stage 3 involves further generalization, limitation, reduction of uncertainty, or corroboration; it is actually never closed, though activity may be slight and publications few.

In research studies, activity in every stage is potentially perpetual, as information will continue to accrue (stage 1 or 2); study proposals will be made (stage 1) and executed (stage 2) for each line of research that might produce results; and corroborating or generalizing studies may be undertaken at any time to increase the validity and reliability of completed research (stage 3).

As one considers how the work of research is done in the VNS stages, the advantages of the RW will become apparent. One of the greatest advantages of a RW is that due to its size, the diversity of interests of the team members will be high, including methodologists, theoreticians and empirical researchers. Staffing levels may reach the point that there is a semi-permanent cadre of trained research assistants that may serve several research projects. Since a RW is organized around a rich issue domain, there can be a stream of research projects passing through, providing research positions for the diverse team. Once operating under full power, the RW will always provide research of interest to any specialty or research orientation.

The RW concept includes several repositories that can store and organize data, information and knowledge both new and old. Research Web Essays can describe the objects and processes of the issue domain. Annotated HyperBibliographies (AHB) can

store references in an interactive annotatable format; and Annotated HyperGlossaries (AHG) can do the same for the vocabulary of the issue domain. Annotatable models make theory accessible to criticism at even the finest level. DocReview is the critical apparatus for most of the documents produced by the research team.

2.3.3.1 Stage 1

Stage 1 activity in a research project includes all activity from convening a research team to the completion of a research plan. Conducting research from within a Research Web provides enormous advantages. Convening a team is easier since members of the RW team are already interested and involved, perhaps only a few new members need be recruited. Gathering information is easier since there is already a knowledge base in place and the means to add new information is present in the RW infrastructure. Preparation of models of the project's focal topic is easier since a mature RW has models of the issue domain already in place.

The team is convened.

The initiators of the research initiative must first gather the core of the research team. What attracts researchers to a team? First, of course the problem itself, then the other team members, funding, and finally, how the team is to operate. The RW, with its infrastructure in place, can be a force in attracting members, especially members committed to collaboration. If the conveners have a paper outlining the nature of the research issue and the proposed products of the collaboration, then that paper can be presented as a Research Web Essay on RW web site. The invited scholars can mount this position paper in DocReview for annotation. This initial dialog not only informs, but also can be instrumental in setting the stage for the direction of the proposal.

The conveners can put other items on the RW that will inform and perhaps attract other scholars from their networks of scholars: personal web pages, and position papers outlining their personal connection with the proposed research issue. This material will show that the conveners are legitimate, have the power to attract funding and can provide an effective environment for collaboration[216]. The personal web pages should include the CV, but should also go beyond that to show how the research issue fits with their ongoing research interests.

Presenting the earliest efforts as a RW can show prospective collaborators that the conveners already have a collaborative environment in place. With a modest effort, all the web site and collaborative software can be put in place and a facilitator recruited. The content will naturally accumulate as conveners and new members present their thoughts in position papers. An Annotated HyperBibliography can be started with the works cited in the convener's position papers and the initial description of the research issue. Similarly, an Annotated HyperGlossary can be populated with the basic vocabulary of the issue domain. The facilitator and graduate assistants can perform these services.

Nothing succeeds like success, and a RW that is up and running is an impressive start. A section of the RW can be devoted to the description of the RW concept and how it will serve the research team. The RW, by its collaborative nature, is a powerful and demonstrable tool and may favorably influence the decision of the granting agency. In the event that the proposal fails, the work remains and can be revised and extended to make a better argument in the future. The RW can also function as a recruiting device to draw in new talent.

Data, information and knowledge are accumulated.

The research team will need to gather all data, information and knowledge relevant to the project's focal topic. While much basic material will naturally accumulate in the RW, what is most important is a plan for acquiring all the needed material. Part of that plan is building repositories for references (AHB), vocabulary (AHG) and organizing the knowledge (RW Essays). The descriptive model accumulates materials in VNS stage 1. Explanatory models may be initiated in stage 1, but will be finally realized late in stage 2. One of the greatest benefits of the RW is that the models, descriptive, explanatory and simulation, apply not only to the research project, but also to the entire issue domain.

Preliminary research hypotheses are developed.

The proposal must present some initial hypothesis in order to argue for funding. The research plan or proposal that the stage 1 team produces can suggest some hypotheses which are certain to be investigated.

The proposal or study plan is written.

The document is the RW's fundamental unit of knowledge representation. A proposal is a formalized document, a genre whose format is specified by the audience. A working copy of the proposal may be mounted on the RW site and made annotatable with DocReview. In that format every member of the team, and administrators from the department or college, may review and annotate the proposal at any time.

Putting the proposal in electronic format suitable for display on the WWW has a number of advantages. If the proposal states that it is available on the WWW, some referees might be inclined to access it on the web, and thus be tempted to explore links that effectively extend the page length restrictions of the proposal. Of course a web presentation can be much more attractive than a black and white laser print. Links to an

Annotated HyperBibliography or Annotated HyperGlossary might serve to demonstrate the effective use of modern technology by the team's conveners.

The RW web site is organized.

The RW web site's working area needs to be organized during the stage one. The organization should be patterned after one or more organizing principle. For research issues that have a strong locational principle, maps can be the organizing principle. Research issues that are dominated by a process should be organized around a diagram of the process. Some research issues are strongly organized by time, such as a history; for these organization by timeline is useful. Some research issues are strongly hierarchical as are some organizations, or classifications of organisms. These organizing principles will eventually become the foundations of the models describing and explaining the issue domain. The standardized infrastructure of the RW web site is discussed below in §3.5.2

It is very important to avoid structuring the RW site around artificial organization. The organizing principles must come from the issue domain, not around the team members or their institutions. If the team is multidisciplinary, the purpose of the RW is to bring the disciplinary material together; the disciplinary material should not divide the issue domain. Artificial boundaries divide and encourage parochialism.

2.3.3.2 Stage 2

Stage 2 is the central stage of the research effort. Knowledge of the issue domain is built by work directed toward three products. These three empirical products are: a set of hypotheses[217], a set of observations[218] and a study design[219]. Hypotheses will, along with the descriptive model, form the basis for the explanatory model. Observations will contribute directly to the descriptive model. Study designs will result in research

experiments that will be used to support the research papers, or perhaps to become research papers on their own.

VNS provides several methods to help the research team insure that the research process produces complete coverage of the project's focal topic. In this stage, the principal method is called the matching process. This process is a collaborative process that compares patterns that emerge from three complementary points of view within the team. The patterns come from studying the issue while progressing along three different paths, discussed below. The congruence of patterns between these paths shows validity of the research by correspondence.

Essays and research papers are written.

Essays may be produced on any topic by the authoring teams or by individuals. Elements of the models, either objects or processes, will be described in *empirical* essays. Elements of the descriptive model may be fully described in essays that will then become targets for criticism by the research team and assembly points for further knowledge about the element. In a similar manner, hypotheses will become part of essays about the elements of the explanatory model. And perhaps when the research process is fairly mature, essays may, by operationalizing the descriptive and explanatory models, display the knowledge about the elements of a simulation model.

As knowledge grows, a set of *synthetic* essays will emerge. These essays will draw their knowledge base from the empirical essays on the elements of the models. They are the basis of published research papers. It is important to retain a working copy of the essay because it remains the definitive research document. If a research study is developed that shows promise for becoming an experimental protocol that might be repeated with other data, then a *methodological* essay will be a very useful, and publishable, product.

Modeling continues.

In stage 2 the descriptive model is largely completed. The explanatory model becomes the focus of theory-building for the research team. The simulation model may be started, though it will come into use only in stage three.

The information repositories are populated.

As the research progresses in stage 2, additional research literature will be identified. The vocabulary of the issue domain will emerge with alternative meanings and glosses. Glosses and alternative meanings will be quite common in interdisciplinary research projects. Experimental protocols, data (both raw data and reduced data), and metadata from the experiments need to be stored. Increasingly, granting agencies are insisting on the sharing of data. The National Science Foundation "expects PIs to share with other researchers, at no more than incremental cost and within a reasonable time, the data, samples, physical collections and other supporting materials created or gathered in the course of the work[220]."

2.2.3.3 Stage 3

> *Rather than spend[ing] valuable (and limited) resources generating new (and equally uncertain) information concerning the same focal problems, we urge researchers to spend more of their future time and effort pursuing robustness analysis as a means for reducing the uncertainty associated with the findings they already have.* --- Brinberg and McGrath[221]

Stage 3 is a mature state where the objectives of the team are to extend, generalize, and explore the limits of the issue domain. Reduction of uncertainty of the stage 2 results is the goal of stage 3 research. Uncertainty is reduced by three mechanisms: replication, convergence, and boundary search. Replication demonstrates the reliability of the results using the same methods. Convergence, or triangulation, uses maximally different

experimental methods to demonstrate the robustness of the findings. Boundary search, or differentiation attempts to show some of the conditions that will produce results not explained by the theory.

Brinberg and McGrath make a strong case for Stage 3 research[222], pointing out that most research reports describe stage 2 studies and leave the difficult work of generalization to "future research." This future research is seldom done because one cannot build a reputation on replication of research. Certainly the value of determining the boundaries of theory and the building of robustness must be examined. Boundaries and nuances can be explored indefinitely, but certainly there are practical limits to the exploration. Senior researchers should design stage 3 studies, but can delegate the more routine work of execution. Given the importance of the socialization of students and junior researchers, some stage 3 studies might be assigned as exercises in legitimate peripheral participation. Certainly a well-conceived and executed series of studies that contribute to the robustness of the original published stage 2 findings would be publishable.

The stage 3 research is likely to add valuable detail to the descriptive and exploratory models. This sharpening of the empirical description of the issue domain and the theory of its operation is very important feedback that will be incorporated in the RW Essays as they are refined beyond the snapshot bounds of the published reports.

The simulation model can incorporate the stage 3 findings of boundaries. This information will improve the behavior of the model, perhaps to the point where the model can be used to develop hypotheses to chart the boundaries of the theory. Since studies are expensive, finding the best places to test will save resources. A well-developed simulation model can support uncertainty analysis and sensitivity analysis. As the goal of stage 3 research is the reduction of uncertainty, the simulation model is an essential prerequisite for the analyses that can accomplish uncertainty reduction.

2.3.4 The Study Paths

A Research Web will have a team that is capable of combining in many research projects, each potentially leading to research papers, as well as the larger collaboration to study the entire issue domain. VNS assumes that scholars prefer to work in one of the three broad areas of empirical description (empiricists), theory construction (theoreticians), or methodological examination (methodologists). VNS suggests that, in order to create the empirical products of the stage, the team engage in collaborations that combine empiricists, theoreticians and methodologists by pairing the interests to proceed along one of three paths, the *experimental*, the *theoretical* or the *empirical*, each leading to a different research product: a study plan, a set of hypotheses, and a set of observations.

As with all facets of the VNS, the flexibility of the schema neither demands nor precludes any form of team configuration, from one member to many. When I speak of a collaborator, I mean a person, or persons, acting in a role. In the social sciences, it is not unusual to have a single person acting in the roles of empiricist, theoretician and methodologist. More commonly, two people will share the three roles, and, of course, several people may share a single role.

2.4 Technology to Support Communication

Communication technologies fall into two pairs of categories: synchronous or asynchronous, and digital or analog (hardcopy). The research team is free to choose the technologies that suit its members. When choices are made, the communicator needs to consider the ability of the communication to be shared and the ability to be indexed and searched. Most of the older analog technologies are difficult to share (telephone calls, letters, conversations) and generally cannot be easily indexed and searched. Digital communication media are easy to share, copy, and search.

Most computer-mediated communication is asynchronous and digital, hence easy to share and search. Synchronous communication media have a special problem: recording.

While analog copies may be made (tapes, CDs, etc.), the recordings are difficult to use and almost impossible to search. Conversations are an important and critical media; those that have content important to the team need to be transcribed to a digital format and summarized for asynchronous use.

2.4.1 Communication Modes

Communication is the glue that holds together all human social activity. In order to understand communication's importance to collaboration, we must investigate the communication process. The object of communication is to transfer a set of symbols, information, from one human mind to another. The elements of a model of communication contain five elements: the source, the transmitter, the channel (with its noise), the receiver, and the destination[223].

Problems in communication begin with the source. Human sources have limitations in their ability to formulate the communication. Does the person who is the source have knowledge of the symbols necessary to formulate the message? Often not. But is communication is necessary, and the source does the best it can. The human source also may consider its target and compose the message in symbols the target is likely to understand.

Transmission of the formulated symbolic message has a set of problems as well. The set of transmitters available to the source may be limited. If the source is mute, it cannot speak. If a person speaks, the symbols may not be well articulated or heavily accented. If the source is not equipped with suitable technology, it cannot write, or keyboard.

The communication channel may not be available, or it may degrade the message with noise. The Research Web depends on the Internet for its communication channel. The Internet has essentially no noise, but is subject to variable rates of delivery and infrequent interruption of service. For speech, the air in the immediate environment is the

communication channel. It is severely degraded with noise in some situations, such as large crowded lecture halls, or at social gatherings. Telephone channels are usually noise free, but are incapable of reproducing high frequency sound, thus persons with voices in higher ranges may be at a disadvantage.

An adequate receiver for the communication channel may not be available at the destination. The RW requires the availability of a WWW-capable workstation, with a display screen capable of rendering high-resolution graphics. FAX transmits only unacceptably low-resolution graphics and requires a special machine. Low-resolution cameras often used for synchronous teleconferencing cannot transmit either text of graphics, only faces or scenes.

The destination has difficulties just as does the source. Disabilities in hearing or in vision limit the ability of the destination from receiving information. Fortunately, the web browsers have the capability of displaying text in a user-selected size consistent with the reader's visual acuity. Language is often a problem, especially with speech. Some people who do not use English as their spoken language much prefer text for communication in English[224], " … foreign scientists occasionally had difficulties understanding oral English and considered text to be preferable, ...". They can reflect on the context of the text message, or consult translating dictionaries.

All these communication problems, as well as many psychological and sociological problems, are of great importance when designing and operating a Research Web collaboration. The systemic effects of synchronicity of asynchronicity dominate the utility of methods of communication. This section examines, with respect to the RW, the systemic advantages and disadvantages of both synchronous and asynchronous communication. Technology also has effects on utility, especially capacity and the ability of existing software to realize the potential of the communications methods.

There are detailed examinations of the application of the major methods of communication within the RW environment.

Gaines has provided a model of the temporal structure of Internet communication[225]. There are four times involved in this model: the origination time (t1)-- the time from conception to becoming available to the recipient; the discovery time (t2) -- the time it takes the recipient to discover the message; the response time (t3) -- time taken to read, understand, prepare and send a response; and finally the response discovery time (t4) -- the time it takes the originator to discover the response. Many of the differences between communication modes are made clear by an analysis in Gaines' terms. This model shows the continuum of temporal behavior between synchronous and asynchronous communication and publication. Clearly reduction of the four temporal technological barriers will improve the efficiency of any communication. Such reductions should be implemented in any design or design revision to the tools of discourse.

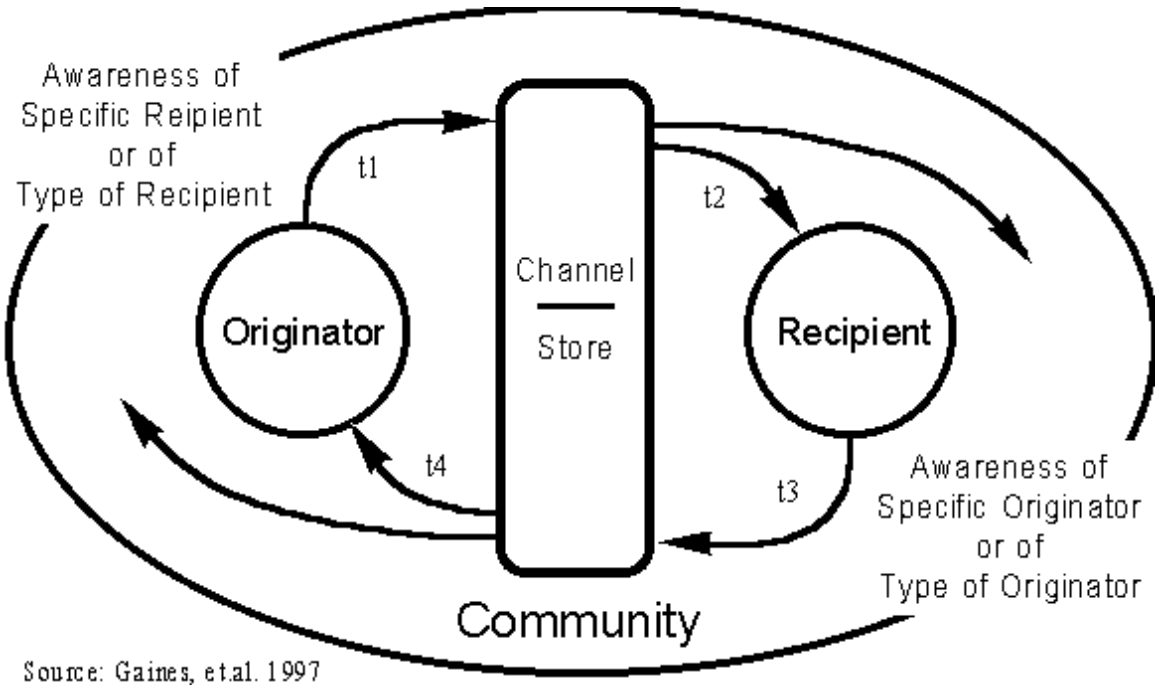


Figure III Punctuated Communication

T1 is the time that elapses between the time that the originator perceives the need to communicate and the time it takes for a message conveying the originator's thoughts to a place where the recipient(s) may receive it. I decompose t1 into three microcomponents: mental processes of the originator (t1a), including formulation of thoughts, reflection on those formulations, and consideration of alternatives; composition of the message (t1b), including selection of words, assembling the argument, and entering the message into the means of communication; and finally reflection, correction and transmission (t1c). T1a is inaccessible to RW technology, except that the RW makes the entire dialog about the issue domain accessible, perhaps assisting in message formulation. T1b can be improved by the use of alternative message entry. Some members may find voice recognition software useful. Use of a modern mailer program with spell checking will improve the quality of the message. T1c can be improved by selecting the fastest means of message

transmission to the pickup point. E-mail can be quite slow compared to WWW page loading, so if the pickup point can be moved to a file that can be accessed by the WWW, T1c can be markedly improved. Many of the RW tools place communications directly into WWW files.

T2 is the time between delivery to the pickup point and the time that the recipient finds it. Bringing the recipient to the pickup point when the message arrives is the problem. Simply waiting for the recipient to call is the slowest method, and may actually prevent receipt of the message. Remember that the e-mail queue is not the only pickup point! Newsgroups are pickup points; the RW has many pickup points for annotations, essentially the web page of the document that is annotated. In the RW, as currently implemented, What'sNew (see §4.7) is a query facility that points the user to all pickup points for messages that have been issued in a certain time period. DocReview, the Annotated HyperBibliography and the Annotated HyperGlossary attempt to shorten the time to pickup by reminding the recipient of activity by e-mail notification. While this is an improvement over simply casting out the messages to the pickup point, things could be better. We are all familiar with "You Have Mail" and beeps when new mail is put in our inboxes; such instant (and hopefully unobtrusive) notification for all RW activity is a future design goal.

T3 and t4 are simply reiterations of t1 and t2. These temporal components are included in the model simply to illustrate "round trip" dialog. The RW does not often have this sort of conversational dialog; in Gaines' terms the RW is World Wide Web publication. t3 and t4 come into play in synchronous dialog.

2.4.1.1 The Time/Place Collaboration Matrix

Communication between humans may be classified in a two-by-two matrix based on time and place. The time values are "same time" and "different time", or synchronous and asynchronous. The place values are "same place" and "different place", or collocated and

distributed. Collocated collaboration is by far the most common form of collaboration in business and classroom, due to the longstanding emphasis on the group in business and educational organizations. Distributed collaboration is a very common form of scientific collaboration, especially in academia.

The asynchronous distributed category contains the communication media most suited to the Research Web concept. While synchronous methods are deprecated, they certainly are not forbidden -- conversation is far too valuable to be discouraged. Collocation exists on several scales: in this work, collocation is considered face-to-face, not just at the same institution or city. Collocation is seldom universal in large teams, especially in academia.

Table III below is a compilation of classifications of communication media from several sources.[226,227,228,229]

Table III Collaboration Modes in Place and Time

	Same Time	Different Time
Same Place	**Conventional Meeting**	**Storyboard Meeting**
	Advantages:	*Advantages:*
	• face-to-face expressions • immediate response • helps building community	• scheduling is easy • respond anytime • leave-behind note
	Disadvantages:	*Disadvantages:*
	• scheduling is difficult • very high cost • power effects • frequent lack of record • isolates those unable to attend	• meeting takes longer
Different Place	**Conference call type meeting, Electronic Conferencing**	**Distributed meeting**
	Advantages:	*Advantages:*
	• no need to travel long distances • immediate response	• scheduling is convenient • no need to travel • submit response anytime • permanent record • reduced power effects
	Disadvantages:	*Disadvantages:*
	• meeting protocols are difficult to interpret to maintain meeting dynamics • limited personal perspective from participants • power effects • Turn-taking rules unclear	• meeting dynamics are different from normal meeting (a new "netiquette") • meeting is never over, lack of milestones • difficult to make decisions • what does non-response mean?

This space/time typology is rather primitive in its treatment of time. Most of the works cited above are discussed in a more technical treatise[230] that considers time in much more detail, using additional terms such as real-time, and concurrent. In their work, they define seven different meanings of synchronous. What is temporally important for our

purposes is whether people must communicate information in scheduled session, or may communicate any time they wish.

Asynchronous distributed collaboration is collaboration among persons in different locations and/or at different times (see Table III above). This mode of collaboration is not popular for several reasons: first much of the technology that enables this sort of collaboration is new, hence not well understood or universally trusted; second, the software necessary to facilitate asynchronous collaboration is still primitive because, unsurprisingly, the intensely market-driven software industry is serving the perceived needs of conventional management by automating current methods rather than innovating; and finally, the current accepted model for collaboration is same time, same place - a well known comfortable, conventional, social activity. Each of the four spatio-temporal modes of collaboration have advantages and disadvantages[231,232,233], these need to be compared and contrasted in order to understand how they might all be brought to support collaboration in research within the issue domain.

This research is directed to the study and development of asynchronous distributed collaboration. My concentration on this mode of collaboration has been influenced by the general neglect of research in asynchronous distributed collaboration in favor of synchronous collocated collaboration, the most familiar mode of collaboration. Business and the academy have honed the tools supporting conventional synchronous collaboration to such a fine point that this form of collaboration has become culturally and economically accepted as the default collaborative behavior.

Realistically, asynchronous distributed collaboration must be facilitated by integration with the Internet, especially the WWW. While the tools employed must be designed to operate on the Internet, the real challenge is to not only facilitate collaboration, but to fully engage the collaborators to such a degree that the tools become genres of communication and information dissemination.[234,235,236,237,238]

Those wishing to categorize the myriad software programs offered to support collaborative applications can utilize a typology that extends the time-place matrix along another dimension: modality[239]. Modality is the combination of communication channels used -- document, audio, and video. This taxonomy is extensible. Nickerson's example includes nine types of collaborative application, but it cannot be universal since software necessarily is targeted toward the functional intentions of the collaborators. DocReview, for instance would be an application type called Document Annotation, a type not found in Nickerson's published typology.

Web-based tools are based on several simple core protocols that have collectively enabled the Internet revolution. These protocols are reliable software commodities that have two major advantages: they are free, and they are stable. Commercial companies are at a competitive disadvantage because they must charge for their products. To create competitive advantages they load minor enhancements on top of the capabilities of the commodities[240]. They are then able to market the products as "improvements" and at the same time undermine the strength of the protocols. Unfortunately, this marketing strategy works. The results are that there are many relatively inexperienced managers and users that accept the "more is better" philosophy. If these people consider how much training time is required to learn how to use these bloated products, perhaps they would see the advantages of simplicity.

Developers can resist these strategies by keeping their products simple and easy to learn. Complexity may be hidden by linking to more advanced features. The facile may become curious and investigate capabilities beyond the basic set. The tools developed for the RW adhere to the concept of hiding complexity and to the solid simple core of existing protocols.

2.4.1.2 Synchronous modes

The most amazing thing about synchronous communication is how *limiting* it is! Every synchronous method revolves about the concept of meeting person to person. Technology has freed us to some extent by allowing us to communicate over distance by telephone and then by television and ultimately, over the Internet. Yet the synchronous methods are, for some purposes, vastly superior to asynchronous means. Fortunately, we understand those purposes well, having used them personally from childhood and, as a society, since speech developed. What are the bases of the frequent superiority of synchronous communication? Rapid negotiation of meaning allows us to plan and learn. Speed of communication allows us to interact rapidly. Ease of communication in face-to-face encounters is undoubtedly the greatest advantage of synchronous communication. For purposes of social bonding, the maintenance of strong ties[241], and ceremony[242], synchronicity and collocation are essential. Temporary person-to-person contact, such as conferences or even phone calls, is often sufficient to maintain ties. Transfer of tacit or poorly expressed knowledge seems to be best accomplished with the strong social ties[243].

Given these advantages, what are the disadvantages? Time, space, power, and permanence. The principal problem is the very practical problem of coming together in time. Scheduling meeting times requires negotiation, unless done by decree. Disruption of personal planning schedules introduces psychological and economic costs. Though being together in time has been made less an onerous task due to the telephone and the Internet, travel is still often necessary to be together in space, be it a walk down the hallway or an intercontinental plane trip. Power effects are present, though often subtle. Synchronous communication always involves turn taking, and turns are controlled by power relationships. Certainly some power effects are positive, especially when decisions must be made. Yet when contributions are sought, power can stifle participation by triggering evaluation anxiety. The recording of synchronous communication is an immense problem. Most synchronous communication is never recorded. Transcriptions of meetings are extremely expensive and often almost

impossible to read. Minutes are seldom accurate and never complete. Audio recordings are difficult to understand without personal knowledge of the speakers. Graphic images from whiteboards are almost never recorded in minutes.

The social effects of synchronous communication on team integrity are very important. Team research depends on unrestricted sharing of information. If the team has a geographically isolated individual, it is important for the better-connected members to keep the isolated member informed about any synchronous communications that have occurred within the synchronously connected group. Failure to recognize the vulnerability of the isolated individual is very likely to result in the withdrawal of that member. In a distributed team with multiperson nodes, meetings at each node seem to be inevitable. Failure to share the proceedings of such meetings destroys the integrity of the team. Meetings are the natural enemy of distributed teams -- they should be avoided whenever possible, and recorded and shared whenever they do occur. The socio-psychological phenomena of *the outsider* and *us versus them* exist to some extent in all human organizations.

2.4.1.3 Asynchronous modes

While synchronous communication is based on speech acts, asynchronous communication is based on message passing. The message is recorded with a device on some medium, is transported by some mechanism, and is then displayed by another device, or in the case of physical messages, often read directly. In most communication, the devices are computer workstations, the medium is electronic and the transport mechanism is the Internet. The address is usually a computer with e-mail software.

Messages can be sent at any time to anyone or in e-mail, to a group known at the destination address. The sender must compose the message carefully since there is no possibility of quickly negotiating the meaning of the message (though the workstation might complain about spelling and grammar). The time it takes to transmit the message

on the Internet is almost negligible. With postal messaging the transmission time is measured in days to weeks. The recipient may pick up the message in seconds, or leave the message idle for weeks. In the RW some messages are posted on the RW's web site and the recipients are notified by e-mail that a message waits. Notification of "messages waiting" and changes in the intellectual content of the RW is an important issue.

Only information can be transmitted on the Internet. Uses remain for asynchronous methods such as postal services to transmit physical items. Voice mail and phone answering machines are useful for very time-sensitive messages. Legal documents requiring signatures are still bound to physical documents, though that is changing rapidly. Large datasets are still transmitted by CD-ROM and tape, but both those applications will soon be shifted to electronic means as technology improves.

The change from atoms to bits is irrevocable and unstoppable.
--- Nicholas Negroponte[244]

2.4.2 Computer-Mediated Communication

Computer-mediated communication (CMC) is communication that has, regardless of source, been transformed into a potentially permanent digital record. All communication in a RW is CMC. Because the contents of the RW are all CMC recorded in digital format, they are permanent and may be recalled at any time. In addition to being permanent, they may also be modified. Modification must, of course, be managed responsibly.

CMC can be divided into two major categories with enormously different properties: synchronous, which takes place nearly simultaneously; and asynchronous, which is stored on receipt by the computer, and served up to the client later, or on request. Gaines has provided a model of the temporal structure of Internet communication[245]. In practice almost all CMC is asynchronous because the inevitable delays caused by transmission, storage and forwarding are disruptive to the smooth flow of communication seen in

synchronous communication. At best, synchronous CMC approaches the jerky character of a long-distance phone call that uses satellite links.

The important differences between synchronous and asynchronous CMC lie not so much in their behavior, but in their psychological, sociological and institutional effects. The RW deprecates synchronous communication because such communication usually goes unrecorded. Similarly it has little use for synchronous CMC because of scheduling difficulties and a decidedly anti-democratic power differential.

It is important to note that the RW does not preclude the use of conventional communication media, synchronous or asynchronous -- it simply insists that the members have the discipline to record those communications in digital format. Meeting minutes and memoranda of communication summarizing important conversations have a very important role to play in the RW. If an important finding is shared between two individuals, the proper way to manage such a message is to send a message to the entire team, thus not only informing every member, but also placing the finding in the permanent searchable message archives.

The principal advantages of asynchronous communication are the temporal decoupling of communicants and the permanent recording of every message. The principal disadvantage is the need to carefully compose the message, an act that takes considerable time when compared to speech. Careful composition is essential because the recipient does not have recourse to rapid verification of his understanding of the terms used and the thrust of the message. The advantage lies with the recipients who are not interrupted by the arrival of synchronous communication, and have time to reflect on the message before replying. The "democratic" nature of e-mail burdens the powerful that now may find it necessary to compose a message on the keyboard instead of using the telephone.

2.4.2.1 Dialog

Dialog is the heart of collaboration. Dialog within a group assumes that there is a common ground, in our context a shared interest in the issue domain of the Research Web. The group attempts to accumulate knowledge by adding to the common ground through contributions. The principal means of adding to the common ground is unilateral action; contributing the right utterance at the right time[246].

Since the beginning of human communication, the building of shared understanding has been enabled by dialog. Before writing, conversation and story telling were the means of sharing knowledge. These forms of dialog are synchronous, and continue in use today. The introduction of writing, and its industrial counterpart, printing, introduced asynchronous dialog. When writing began to record knowledge, the first recorded dialogues were marginal notes to religious texts, and recorded dialogs and debates. In modern times, technology has introduced new media for both synchronous and asynchronous dialog.

Modern synchronous methods basically augment speech while asynchronous methods augment writing. Dialog in either of these families of media depends on interactivity. While interactivity in speech is as old as the human race, interactivity with asynchronous methods is relatively new, beginning with message passing and industrialized with the initiation of modern postal services. Synchronous methods of communication have been a constant target of technology and have recently been epitomized by broadband teleconferencing. The introduction of the Internet and services such as the WWW has produced a vast improvement in asynchronous methods, now providing speed that rivals synchronous methods at near-vanishing monetary cost.

We examine now the different natures of dialog in synchronous and asynchronous methods. Natural dialog, or face-to-face speech in pairs or small groups, is characterized by several social conventions: presentation, acceptance, evidence of understanding, and

turn taking[247]. We can differentiate synchronous and asynchronous dialog by taking a close look at each of these conventions.

Presentation is the initial contribution by one of the team members. In speech, this is usually a sentence and seldom more than a paragraph, a set of a few related sentences. In the RW, presentation is usually longer, from an e-mail message of a few sentences to a complete essay of several thousand words. The reason for such a vast range of sizes is that the contributions of the team are submitted *en Toto* as a complete argument for the consideration of the team members. This mode of contribution is necessary because piecewise acceptance is impossible in asynchronous communication. Also speech is generally conducted in pairs or small groups while asynchronous communication in the RW is always a one or few-to-many transaction. Dialog in very large cognitive chunks can best be effective when the presentation is permanently recorded so working memory is not overwhelmed.

Presentation in speech is impaired by clarity of speech, ambient noise, and poor hearing[248]. From a theoretical communication standpoint much of speech and written communication share equally in problems of symbol selection and interpretation[249] (encoding, decoding) since words are symbols common to both. Written presentations have a much lower susceptibility to errors of transmission. While speech is aided by inflection and gesture, writing may also be inflected with alternative character fonts or underlining. Since speech is ephemeral, understanding diminishes with time; writing does not have that problem as the reader can return to reflect on the words. The issue of size of the contribution presented is more difficult. Speech is a serial process, while presentation of a large written contribution is a parallel process that presents several propositions at the same time. Readers can select from these propositions in a parallel manner by screening and then evaluate the selected proposition in a serial manner by reading.

Acceptance is the process of ensuring understanding between communicants. In speech, the turnaround time is very short, so a constant stream of presentations and acceptances can be interchanged efficiently. Among communicants known to each other, gestures often serve effectively as acceptance. Should acceptance be refused, repair transactions will be initiated immediately. In asynchronous communication, gestures are absent and turnaround time is highly variable and often long, so alternation of presentation and acceptance is impossible. Asynchronous collaboration in the RW depends on three mechanisms to monitor acceptance of a contribution: first, implied acceptance; second, the existence of a common language; and finally, responsibility of members to review contributions.

Acceptance in asynchronous presentation is implied by silence. Note that there is no conditional acceptance by silence; any response, other than praise, is a request to modify or clarify the presentation. Implicit acceptance is, however, not unequivocal because people are usually reluctant to challenge statements unless they are easily refuted. When it is important to be certain of approval, poll the team[250].

The existence of a common language is assumed in both speech and writing; but the RW has the advantage of having an integral glossary of terms. In a series of speech dialogs the meaning of a term can drift, because the context of each meeting is different. The asynchronous dialog in the RW has an interactive glossary so the team can have a recorded dialog about any term.

Reviewing presentations is required both in speech and in asynchronous presentation, but this process is subject to many more damaging power effects in speech. Review (as acceptance) requires evaluation, criticism, and sometimes the admission of ignorance. Face-to-face criticism, especially in a group situation is difficult for many people. It is widely held that power effects are diminished in asynchronous communication[251,252,253,254]. In an asynchronous situation, the reviewer has the

opportunity to communicate off the record with peers or with the presenter, thus avoiding embarrassment. Review responsibility is implied in speech, but is more explicit in review of asynchronous contributions. Hiding disinterest behind nods of acceptance is common in speech; but its equivalent in asynchronous dialog, silence, is almost an admission of inattention. This is so because in large asynchronous contributions, there is obviously little likelihood of total concurrence.

In speech the commonest form of contribution is turn taking. Each turn is a collective act consisting of two participatory phases: presentation and acceptance. The median length of these contributions is 9 to 13 words[255]. Each turn adds to the common ground. Obviously, turns must be taken in a synchronous mode with very short pauses between phases. Turn taking in this sense is clearly impossible in an asynchronous environment.

So, what is the equivalent to a turn in the asynchronous environment of the Research Web? Another speech act is the assertion, which is an autonomous act, not a participatory act. Assertions, *unless rejected*, become part of the common ground. Thus in the asynchronous environment, turns consist of assertions (autonomous contributions) in written form, and the frequently lengthy exchanges of commentary (participatory contributions) required to repair the original assertion. In the Research Web, most contributions are understood to be plastic, and each contribution, or work object, has an owner who is responsible for making repairs based on commentary. Contributions in the RW are all autonomous, and vary in size from a complete RW Essay down to a brief comment attached to an assertion.

The apparently different natures of synchronous and asynchronous dialog are resolved by having a common end: adding to the common ground. Synchronous dialog quickly adds to common ground in very small units. Asynchronous dialog adds to the common ground either rapidly or slowly depending on the size of the assertion. Adding a reference to the Annotated HyperBibliography is unlikely to require repair, so adds to the common

ground quickly; on the other hand, asynchronous dialog on a large RW Essay may last the life of the RW effort.

2.4.2.2 Criticism

> *"... Critiques pose problems and puzzles that have the potential to stimulate pathbreaking research. Critiques suggest novel directions and let newcomers know that there is a place for them in the solution of current problems. Critiques keep researchers from becoming complacent. ... The best way to advance knowledge, it follows, is to foment a constant stream of criticism and response. ..."*
>
> --- Marshall Scott Poole[256]

Critical annotation of documents is a practice fundamental to scholarly activity. Popper has argued that criticism is the basis of rationality in science[257]. Miller, in his book on critical rationalism, suggests that all science be accepted, if falsifiable, until it is proven false[258]. Justificationism, on the other hand, suggests that the better approach is to justify the hypotheses before the science is accepted. This is the approach taken by the RW, that criticism is a positive force for construction of science from the very beginning of work.

Critical social theory [of Horkheimer, Adorno, Fromm, and Marcuse] states that reason and critique are inseparable, and that researchers using this approach must collaborate with those affected by the research and must open the research to public debate and critical reflection[259]. The RW fulfills these requirements and also opens the way to action science through participative design and execution. Action research has been practiced with the RW's tools through the collaborative development of a metadata collection system[260].

Criticism can take substantial forms such as disagreement, clarifications, alternative explanations, and addition of information; or trivial forms such as grammatical or orthographical corrections. In the RW, all knowledge presented is in the form of a document, and the major documents all have a critical apparatus, or means of criticizing. RW Essays and other lengthy documents such as position papers and informative introductions may be presented in DocReview (see §4.3). The Annotated

HyperBibliography and Annotated HyperGlossary have built in annotation capabilities for each entry.

The environment that encompasses the Research Web and the team must support the concepts of fairness and competence in order to ensure that criticism can operate properly. The Ideal Speech Situation (see Table IV, below) described by Habermas as modified by Webler fulfills the conditions of fairness and competence required for the RW[261]. Webler's rules are designed for public participation and are therefore designed for a less elite membership. See the author's notes for application to the RW environment.

Table IV Criticism Insured by the Competent Ideal Speech Situation

Rule	Fairness	Competence
1. Every potential discourse participant must meet minimal societal standards for cognitive and linguistic competence	Anyone may participate (a)	Minimal standards for cognitive and lingual competence (b)
2. Every discourse participant must have access to the knowledge needed to make validity claims and criticize others'.	Assert validity claims (c)	Access to the knowledge (d)
3. Speakers must verify the results of any attempt to translate expressive claims.	Challenge validity claims (e)	Consensually-approved translation scheme (f)
4. Judgments about conflicting validity claims must be made using the most reliable methodological techniques available.	Influence final determinations of validity (g)	Most reliable methodological techniques available (h)

--- from Webler 1995[262]

Author's notes applying the situation to the Research Web environment
(a) Anyone on the research team plus invited guests.
(b) Standards must be relaxed to allow participation of junior members.
(c) A major responsibility. If research is viewed as argumentation, then validity must be established.
(d) Every part of the team RW must be open (exception is authoring team working areas). Tacit knowledge should be made explicit.
(e) See (c).
(f) Members may query any other member to clarify a document.
(g) This condition is met by leaving all documents open to criticism.
(h) Support software should be improved by participatory design.

Each member of the research team has the duty to review and criticize every document that lies within the member's area of interest. The member is expected to be responsible for the accuracy of the commentary, and should provide references to literature where known. Criticism should be attributed to the critic as a matter of responsibility. In asynchronous dialog, presentations are assertions; and criticism of assertions is the participatory mechanism that provides information to repair the presentation. The owner of the presentation is charged with editing the presentation from time to time and representing it as a new assertion. This cycle of presentation, criticism, editing and representation is the engine of refinement of the assertions that constitute the common ground of the research team.

2.5 A Conceptual Framework for Research Collaboration

The conceptual framework (see figure IV, below) has three mutually supporting legs; the first is the philosophy of realism, which provides the backing for the use of models in theory building. The second leg is the methodology of the VNS that, with its tripartite domain organization, provides a practical locus for the models proposed by realism. The third leg of the conceptual framework is the Research Web, the interactive environment that provides a home for models that realism proposes, and the support for the activities of the three domains of the VNS methodology. The Research Web, examined in detail in Chapter 4, is the repository for all team documents and the knowledgebase that the team assembles. The Research Web enables several fundamental functions in document management including criticism and successive refinement.

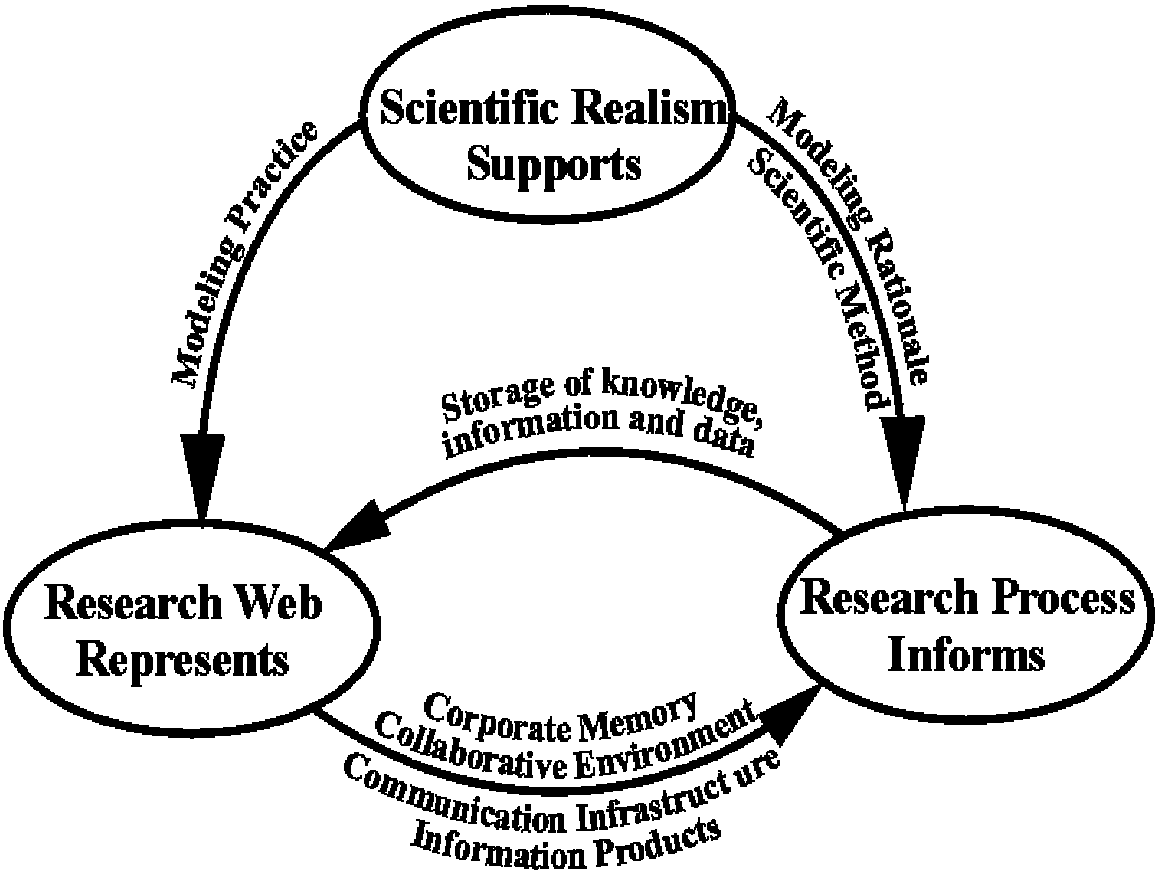


Figure IV Conceptual Framework

Notes to Chapter 2

[1] Cartwright, Shomar and Suárez 1995, 140
[2] Aronson, Harrè, and Way 1995
[3] Denzau and North 1994, 3
[4] Sayer 1992, 104
[5] Ziegler 1990, 2, 27
[6] DeMarco 1979
[7] *ibid.*
[8] McMinamin and Palmer 1984
[9] Rumbaugh 1991
[10] Ziegler 1990
[11] Booch, Rumbaugh and Jacobson 1999
[12] Fowler 1997
[13] Halpin 1999
[14] Dunn 1982
[15] Harrè 1970, Chapt.1
[16] Denzau and North 1994, 13
[17] Zander and Kogut 1995, 77
[18] Augier and Vendelø 1999, 255
[19] Aronson, Harrè, and Way 1995, 51
[20] Harrè 1978, 275
[21] Blalock 1968, 24
[22] Blalock 1984, 57
[23] Blalock 1990, 34
[24] Hox and Mellenbergh 1990, 124
[25] Blalock 1990, 35
[26] von Bertalanffey 1968
[27] ibid. 27
[28] Morrill 1987, 540
[29] Eliasmith, 1999
[30] Shank 1998, 843
[31] *ibid.,* 848
[32] *ibid.,* 848
[33] Shank and Cunningham 1996
[34] DeMarco 1978
[35] Ziman 1976, Chapter 5
[36] Weimer 1979, 78
[37] Price 1963
[38] Veaner 1985, 6
[39] Latour 1987, 33+
[40] Garfield 1979, chapt. 10
[41] Taubes 1993
[42] Tomney and Burton 1998
[43] Holoviak 2001, 14
[44] Cronin and McKim 1996, 169
[45] van Raan 1997, 448
[46] Odlyzko 1995, 84
[47] Lawrence 2001

[48] von Bertalanffy 1968, Chapter 6
[49] Brinberg and McGrath 1985, 160
[50] Sterman 1991, 219
[51] Leon Lederman 1969
[52] Merton 1973
[53] Merton 1973
[54] Cohen 1995, 1706
[55] Beaver and Rosen 1978, 68
[56] Beaver and Rosen 1978
[57] van Raan 1997, 446
[58] Latour and Woolgar 1979, 200 *et. seq.*
[59] Rice 1991, 12
[60] Boyer 1990, 36
[61] Star and Ruhleder 1996, 126
[62] Birchard 2002
[63] Langston 1996
[64] Lindsay 1978
[65] Cronin, *et.al*. 1998
[66] Foote 1999, 115
[67] Matzat 2001, 220
[68] Lave and Wenger 1991
[69] Cronin 1995
[70] Trice and Beyer 1984
[71] Glenwick and Burka 1978, 213
[72] Daniels 1990
[73] Hardin 1998, 8
[74] Cronin 1995
[75] Austin and Baldwin 1991, 23
[76] Endersby 1996
[77] AAG 1993
[78] Liff 1998
[79] Etzioni and Etzioni 1999
[80] Fehr and Gächter 2000, 990
[81] Sanderson 1996
[82] Cohen 2000
[83] Leatherdale 1974
[84] Thagard 1993
[85] Osgood 1953
[86] McKendree and Mayes 1997
[87] Shrage 2000, 206
[88] Gächer and Fehr 1999, 343
[89] *ibid.*, 362
[90] Sanderson 1996, 96
[91] Younglove-Webb *et.al*. 1999
[92] Damrosch 1995, 195
[93] T.W. Malone and K. Crowston 1994, 90
[94] Schrage 1990
[95] Austin and Baldwin 1991, 4
[96] Henderson 1991, 452
[97] Oliver, Marwell, and Teixeira 1985
[98] Markus 1987

[99] McGovern and Norton 2002, 107
[100] Crow, Levine and Nager 1992, 739
[101] Weinberger 1999
[102] Cronin 1995
[103] Hutchins 1991
[104] Maienschein 1993
[105] Webler 1995, 59
[106] Pröpper 1993, 82
[107] Eisenhart and Borko 1993, 93
[108] Pröpper 1993, 82
[109] Endersby 1993, 377)
[110] Maienschein 1993
[111] Nyerges 1999, personal comm.
[112] Maienschein 1993, 180
[113] Kraut, Galegher, Egido 1988, 35
[114] Endersby 1996, 377
[115] Poole 1994
[116] Endersby 1996, 377
[117] Lave and Wenger 1991
[118] Austin and Baldwin 1991
[119] Shreeve, *et.al.* 1986
[120] Thagard 1997, §4.3
[121] Austin and Baldwin 1991
[122] Thagard 1997, §4.4
[123] Maienschein 1993, 178
[124] Thagard 1997, §4.2
[125] Thagard 1997, §4.1
[126] Endersby 1996, 389
[127] Austin and Baldwin 1991, 23
[128] Austin and Baldwin 1991, 22
[129] Schön 1983, 202
[130] Funtowicz and Ravetz 1993, 740
[131] McKay and Marshall 2001
[132] Wadsworth 1998
[133] Ruhleder and King 1991, 342
[134] Miller 1995
[135] Marshall Scott Poole 1994
[136] Ngwenyama 1991, 269
[137] Maienschein 1993
[138] NSF 1999
[139] NATO 2001
[140] Maienschein 1993, 177
[141] Wood and Gray 1991, 152
[142] Chrisman 1987, 1369
[143] Tierney and Bensimon 1996
[144] Bohen and Stiles 1998, 46
[145] Foote 1999, 115
[146] *ibid.,* 116
[147] Cronin 1995
[148] Dubrovski, Kiesler and Sethna 1991
[149] Rice 1991, 8

[150] Baldwin and Austin 1995, 64
[151] Stark 1995, 181
[152] Cronin, 1995
[153] Ruhleder 1995, 50
[154] Bohen and Stiles 1998, 43
[155] Isserman 2000, 311
[156] Wasow 1992, 486
[157] Tierney and Bensimon 1996, 66
[158] Woit 2002, 112
[159] Baldwin and Austin 1995, 65
[160] Foote 1999, 115
[161] Bohen and Stiles 1998, 46
[162] Kahn and Prager 1994
[163] Stark 1995, 181
[164] Dinges 1977, 139
[165] *ibid.*, 139
[166] Crow, Levine and Nager 1992
[167] Stark 1995,182
[168] Poole 1994, 26
[169] Kennedy 1997, 148
[170] Hutchens 1998, 36
[171] Kahn and Prager 1994
[172] Bohen and Stiles 1998, 42
[173] Tierney and Bensimon 1996, 38
[174] Kahn 1993, 5
[175] Damrosch 1995
[176] Damrosch 1995, 190
[177] Ervin and Fox 1994
[178] Thagard 1997, §4.2
[179] Citera *et.al.* 1995, 556
[180] Brody 1996, 49
[181] Kahn and Prager 1994
[182] Diamond 2001, 24
[183] Geert Hofstede 1991
[184] Hofstede and Bond 1984
[185] Tsou 1998
[186] Hsü 1992, 88-94
[187] Hébert 1986, 15
[188] Harvey 1997, 144
[189] Bantz 1993, 11 *et.seq*.
[190] Hannerz 1992
[191] Heaton 1998, 214
[192] Kurland and Egan 1996, 398
[193] Bantz 1993, 12
[194] *ibid.,* 13
[195] *ibid.,* 16
[196] Dubrovski, Kiesler and Sethna 1991, 121
[197] Bantz 1993, 11
[198] Bantz 1992, 14
[199] DeMente 1994
[200] Feghali 1997, 158-159

[201] Brinberg and McGrath 1985
[202] Carstensen 2000
[203] Younglove-Webb *et.al.* 1999, 426
[204] Bero and Rennie 1995
[205] Markus 1987, 494
[206] Brinberg and McGrath, 1985, 33
[207] Jankowski and Nyerges 2001, Chapter 4
[208] Brinberg and McGrath 1985, 30
[209] Funtowicz and Ravetz 1993, 740
[210] Donald Schön 1982
[211] Nyerges and Jankowski 1997
[212] Jankowski and Nyerges 2001
[213] Brinberg and McGrath 1985, 38
[214] Nyerges, Jankowski and Drew 2002, 10
[215] Blalock 1990
[216] Wood and Gray 1991
[217] Brinberg and McGrath 1985, 64
[218] *ibid*. p66
[219] *ibid*. p62
[220] NSF 1999, Chapter VII.H
[221] Brinberg and McGrath 1985, 136
[222] Brinberg and McGrath 1985, Chapter 5
[223] Osgood 1954, 1
[224] Sanderson 1996, 107
[225] Gaines, *et.al*.1997, 1000
[226] Ellis, Gibbs and Rein 1991
[227] Johansen 1992
[228] Nyerges 1995, 269
[229] Hansen, *et.al*. 1999
[230] Antillanca and Fuller 1999
[231] Ellis, Gibbs and Rein 1991
[232] Johansen 1992
[233] Nyerges 1995, 269
[234] Bazerman 1988
[235] Orlikowsi and Yates 1994
[236] Agre 1995
[237] Oravec 1996, Chapter 1
[238] Orlikowsi and Yates 1998
[239] Nickerson 1997
[240] Glass 2001
[241] Wellman 2000
[242] Trice and Beyer 1984
[243] Augier and Vendelø 1999, 255
[244] Nicholas Negroponte 1995, 4
[245] Gaines, *et.al.* 1997, 1000
[246] Clark and Schaefer 1992
[247] Clark and Schaefer 1992, 164
[248] Shannon and Weaver 1949
[249] Osgood 1954
[250] Sheard 1995
[251] Dubrovski, Kiesler, and Sethna 1991,139

[252] Spears and Lea 1994, 448
[253] Kurland and Egan 1996
[254] Postmes, Spears and Lea 1998, 693
[255] Clark and Schaefer 1992,174
[256] Poole 1994, 21
[257] Weimar 1979, 81
[258] Miller 1994, 6
[259] Ngwenyama 1991, 269
[260] Hendricksen 1998d
[261] Webler 1995, 59
[262] Webler 1995, 57-58