

Chapter Three: Methodology

The reason why we are on a higher imaginative level is not because we have finer imaginations, but because we have better instruments. In science, the most important thing that has happened over the last forty years is the advance in instrumental design (Whitehead, 1925).

3.1 Overview of Study

For this study, I chose an unobtrusive research design--a content analysis, to critically explore whether culture effects participation of distributed participants when they collaborate through email. For the source data, I used archived emails drawn from the participation of Civil Society participants in the World Summit on Information Society (WSIS). Within the archival email messages, I explored the cultural variations using a theoretical lens called high context and low context. Additionally, I drew upon existing literature in this domain to provide context and analysis. As a result, this study yielded rich descriptions of the ways participants of these two cultural orientations participated and contributed to the different decision-making stages such as problem identification, response to ideas and deliberation, proposal making and generation of ideas, and solution. At each stage, I identified patterns of behaviors from high context and low context cultures. Given the literature review and theoretical framework presented in Chapter Two, the grand tour research question (RQ) that this study aimed to answer was:

Is there evidence of an influence of high-context and low-context cultural communication styles on WSIS Civil Society decision making processes in globally distributed collaboration using email?

Specifically:

- RQ1. How does the globally distributed WSIS Civil Society decision making processes compare with Adler's and Kingdon's proposed models?
- RQ2. Are there discernable patterns of variations evident in the email messages?
 - (a) If so, is Hall's high context vs. low context cultural communication styles sufficient to explain these variations?
 - (b) If not, what other factors might also be operating?
- RQ3. What is the effect of cultural communication style on individual contributions in the globally distributed WSIS Civil Society decision-making processes?

To answer the research questions as proposed, the study was carried out in several stages (see Figure 3.1). In Stage One, I developed the research design for the empirical study based on two preliminary studies that I had conducted. These studies used Q-sort and interview methods to examine the effect of culture on communication styles and the use of computer-mediated communication technology. The studies also formed groundwork for understanding the construct, theoretical lens, phenomenon, and methodological issues. Equipped with this foundation, in Stage Two, I analyzed a small subset of the source data which comprised three and half months of archival messages.

The purpose of this phase was to gain some sense of the nature of the data, to provide some direction on how best to carry out the study, and to test the coding scheme. The data were sampled from the complete archival datasets that were subsequently used in this study. The pilot dataset also offered a chance to train the coders and to get them familiar with Atlas TI, the computer aided qualitative data analysis software (CAQDAS) selected for the study. In the final stage, I created an analytic framework with three different analyses (situational analysis, cultural analysis and decision-making analysis) with the full set of data covering six months. In this stage, I also conducted inter coder reliability checks to ensure a valid and reliable result.

The rest of this chapter is divided into four sections:

- Methodological background — discussing the research context and literature review;
- Empirical study (Stage One)— describing research designs of content analysis, data collection procedures, and data analytic framework;
- Analysis on pilot data (Stage Two) — describing the process of using a sample data to work on, training the coders, testing the created coding scheme, revision of coding procedures, and refinement of coding scheme;
- Data analysis (Stage Three) — describing the process of achieving high inter coder agreement percentage, and addressing the issue of validity and reliability.

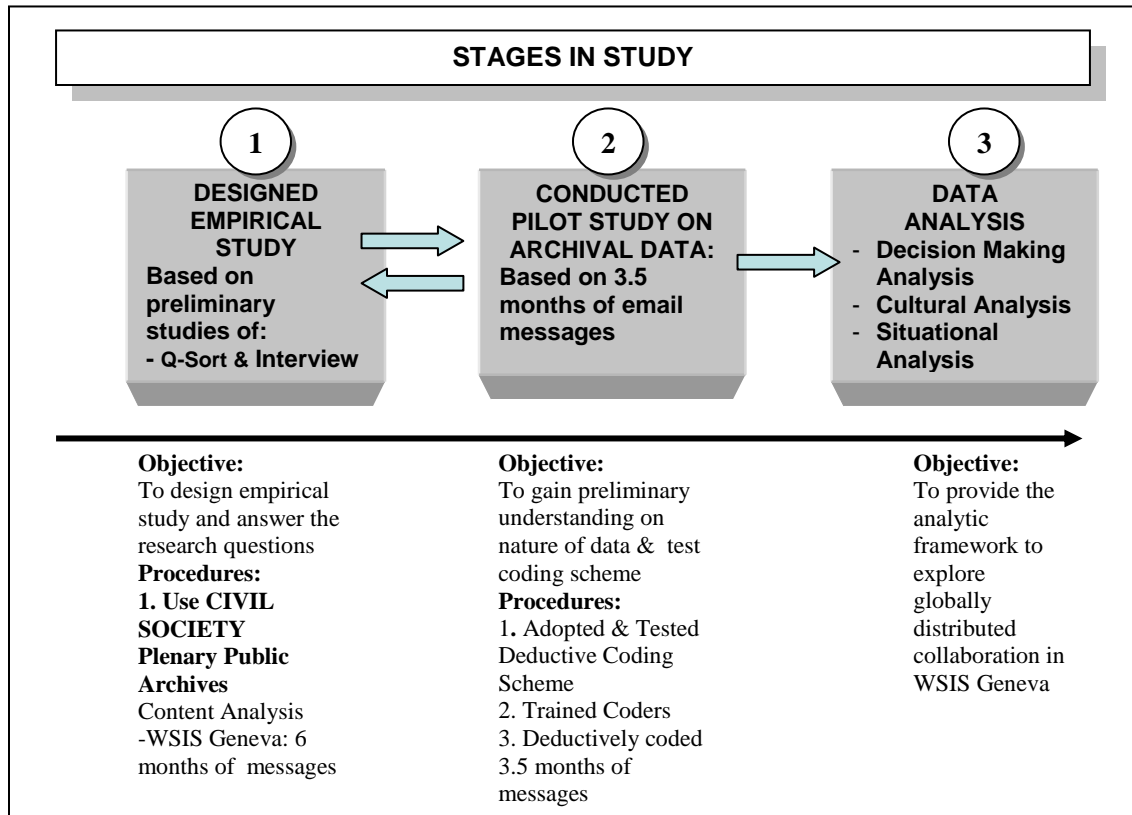


Figure 3.1: Steps Implemented in the Empirical Study

3.2 The Setting: World Summit on the Information Society (WSIS)

WSIS as an international conference was notably different from any other UN conference because the summit took place in two different phases—Phase I in Geneva (December 10-12, 2003) and phase II in Tunis (November 2005). Although the actual conference itself occurred over only a few days, but the preparatory and follow-up processes occurred over a period of several years.

Klein (2005) noted that the activities were most intense during the preparatory phase. Between 2001 and 2003, there were two series of meetings: preparatory meetings (prepcoms) and regional meetings. All three prepcoms were held in Geneva at 6-month intervals. The regional meetings were carried out in various locations around the world but within a shorter time frame. All the meetings were designed to gather input from around the world and to prepare the documents that were to be finalized and adopted as an outcome of WSIS Geneva (see Figure 3.2).

3.2.1 Participation of Civil Society Participants Using Email

This study makes two assumptions regarding email participation of WSIS Civil Society participants. First, all participants who were registered on the email listservs had access to the technology, making the issue of digital divide irrelevant in this research context. Second, all participants freely made the choice to use a lean medium of communication, specifically email. This last assumption is supported by empirical findings that 89% of all civil society participants used email listservs in preference to other CMC advanced tools such as blogs or wiki webs (Cogburn, 2005).

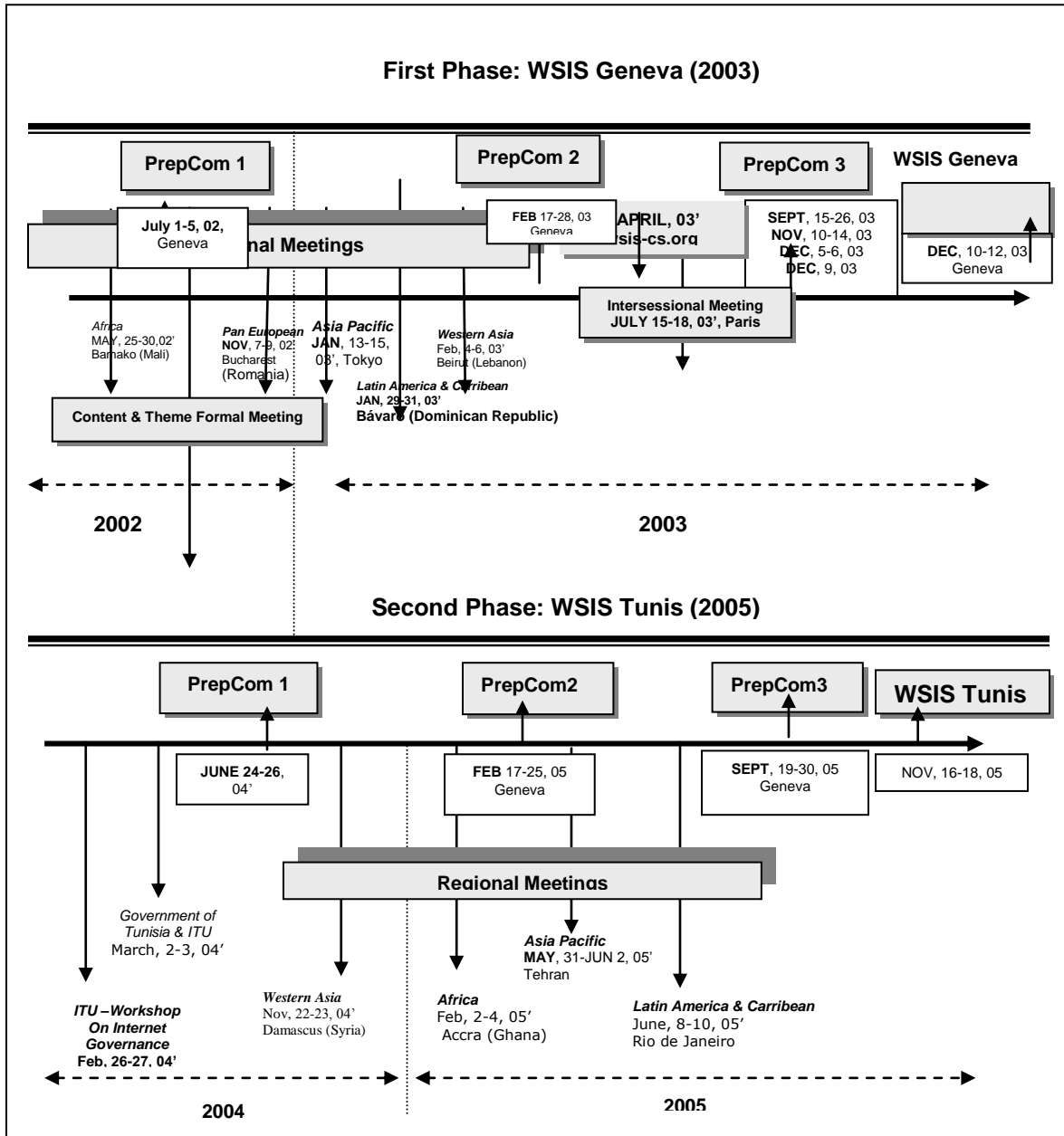


Figure 3.2: Timeline for the Two-Phased WSIS (Geneva, December 10-12, 2003 and Tunis, November 16-18, 2005)

3.3 Outline of Study

3.3.1 Overview of Content Analysis Design

The purpose of this study was to understand how culture affects civil society participation when participants use email to collaborate. The study looked at a specific aspect of culture called context, which takes two forms: high context or low context. The medium of collaboration for participation was email. The participation of Civil Society was analyzed based on their contributions to decision-making in these areas: (1) problem recognition/agenda setting, (2) information search, (3) construction of alternatives, and (4) choice. An individual may contribute differently depending on whether he/she is high context or low context.

3.3.2 Data Collection Procedures

3.3.2.1 Archival Email Messages

Figure 3.3 presents the general structure of WSIS Geneva in order to give an overall perspective on and a broader understanding of the conference, and shows where and how Civil Society participation fits into the larger process. It is important to note that WSIS was unique among other UN conferences because UN actively includes the participation of three groups of stakeholders: government, public sector, and Civil Society in its deliberations. Although the governmental bodies would make the final decisions in WSIS Geneva and Tunis, the three groups' contributions were incorporated in those decisions. In this study, I have chosen to focus on one group of stakeholders – the Civil Society.

The Civil Society participants used email in the form of a listserv as their primary vehicle for communication and collaboration. The participants came from all parts of the world – Asia, Europe, Africa, North America, and South America. With the capability of email they did not need to meet in face-to-face encounters, though many had the opportunity to do so. Interestingly, the email listserv generated many successful negotiations and exerted a strong influence on the process. Many email listservs were created and organized by Civil Society (CS) including CS Plenary, CS Content and Themes, CS Bureau, Thematic Caucuses and Working groups such as Cultural and Linguistic Diversity, Gender, Youth, Media, Person Disabilities, and many more (www.wsis-cs.org). Although this study provides a brief general overview of the entire WSIS event in the following section, this study's focus is on the most important, prestigious, and largest listserv, CS Plenary, because it comprised the most data to understand the collaborative behaviors from a cultural stand point. Also this listserv was the main medium that integrates participants from other groups or caucuses to work together to discuss problems and agenda ranging from issues centered on infrastructure, resources, language, to Civil Society structures and many more.

As noted in previous chapters, WSIS was a two-phase process – WSIS Geneva in December 2003, and WSIS Tunis in November 2005. In total there were 31 months of distributed collaborations via the email listservs, from April 2003 to November 2005. However, preliminary WSIS efforts began as early as May 2002, in the form of preparatory meetings for the WSIS in Geneva, resulting to 42 months of collaborative efforts within Civil Society participants. The phases each had distinct outcomes, goals, and objectives. The first phase of WSIS produced two documents: the Declaration of Principles and Plan of Action.

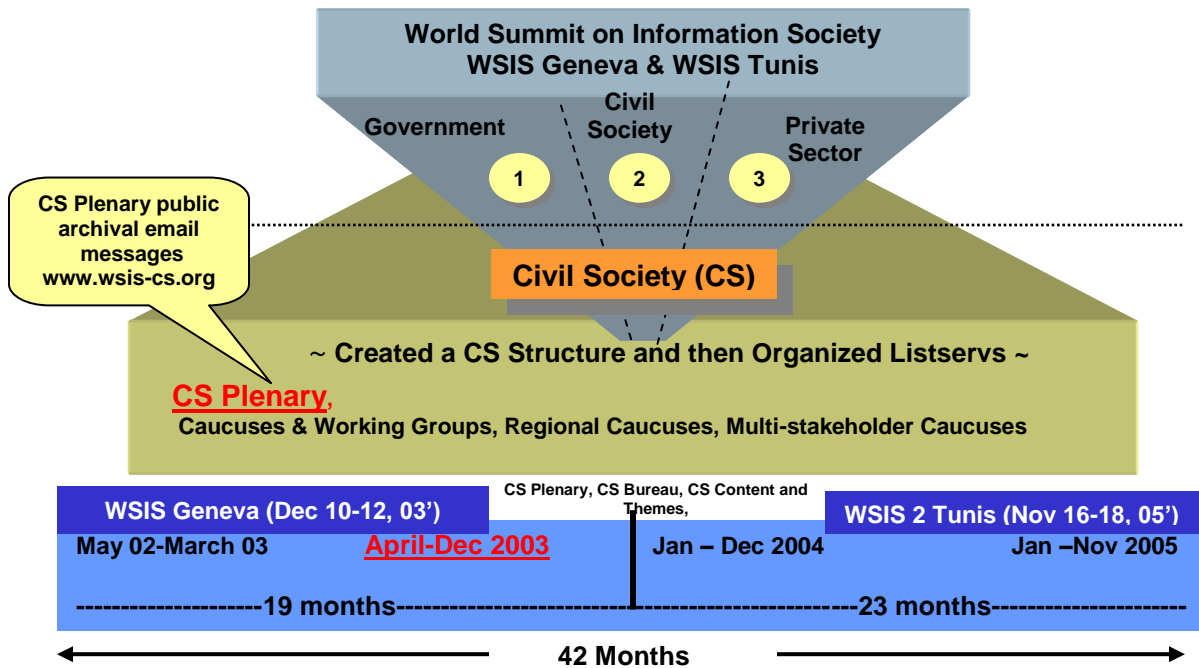


Figure 3.3: General Framework of WSIS

The second phase of WSIS focused on implementation and strategies. To fully understand the dynamic of the distributed collaborations, one must take into account the many ongoing face-to-face conferences, regional meetings, and preparatory meetings (see Figure 3.2). Although some data for these interim meetings, such as the minutes of what transpired, was available, it was not analyzed in this study with the exception of using the dates of the face-to-face meetings as punctuations whereby to observe the patterns of email participation.

This study chose to analyze WSIS Geneva because that phase most closely reflected the dynamics of human collaboration in a globally distributed environment. The primary goal for Civil Society participants for this phase was to effectively influence the language¹ or content of the two documents (Declaration of Principles and Plan of Action) including questions such as what principles were to be accepted and included, what problematic areas should be looked at for further action plans, and so on. Serious and complex decisions had to be made during this process; hence, the challenges of intercultural communication faced by Civil Society participants during email discussions were apparent in the decision-making process. On the other hand, in WSIS Tunis, the dynamics of globally distributed collaboration were more complicated and diffuse since the focus and effort was no longer on just two documents. It would thus have been far more difficult to isolate and examine the decision-making process using data from this second phase of WSIS.

As shown in Figure 3.3, this study focused on Civil Society participation in the plenary listserv in WSIS Geneva. The analysis was conducted on data from a six-month period (July to December 2003) because these were the most active months in terms of observable online communication behaviors. In the early months of the email initiations (April-June 2003), Civil Society participants did not contribute substantially to the email discussions because there was nothing much going on in WSIS or the preparatory meetings outside the virtual plenary listserv.

The CSP structure in WSIS approved a Virtual CS Plenary (VCSP) as an electronic forum for use between physical meetings of the CSP, to promote debate and greater transparency in the organization of civil society during the Summit preparatory process. The listservs were available to all

¹ 'Language' as used here and by the listserv members refers to the intellectual content of the two primary documents. It does not pertain to English, French, Spanish, etc. although the problem of translation was also discussed at length on the listserv.

accredited bodies, to registered bodies at WSIS meetings, and to all civil society entities with an active interest in the WSIS. The archival email messages from the discussions among CSP team members serve as the primary source data for this study. The message content covers a broad range of topics, from administrative information to technical issues, from questions of translation to specific wording of sentences in the two documents.

3.3.2.2 Cleaning the Dataset

The data went through several cleaning procedures². First, the data sets were downloaded from the raw archival data available from the Civil Society Plenary public archives. The archives were in an MBOX file, a readable text-based mail storage format. The data were imported into the Mozilla Thunderbird email client (Thunderbird was chosen because it is open-source software, and numerous extensions or plug-ins). Importing as email preserved the archival dataset's organization message by message, much like the original sender and recipient would have seen. Additionally, the messages were much easier to read in email form since the email client provided a general idea of the magnitude and amount of data.

Once the raw data were downloaded, the data set was reviewed using Ultra-Edit, a text-editor capable of dealing with very large text files (the largest file in the data set was 110MB). Ultra-Edit also displays the messages with line numbers which eases the reading of very large files. Messages were removed³ if it contained any of the following: (1) purely HTML codes, (2) purely attachments. In all emails, the message headers were retained but only some complex information of the headers that provide technical information of the senders were removed. Finally, the cleaned dataset was organized by month and then each message was saved as a text file. With all these steps taken, the data were then imported into Atlas Ti Version 5.0, a qualitative data analysis software for analysis.

3.3.3 Data Analysis Procedures

In this study, I focused on effective participation only in three stages of decision-making: (1) problem identification, (2) proposal making including idea generations, giving suggestions or expressing opinions, and (3) solution. Initially, this study identified four stages based on Kingdon's (1995) and Adler's (1997) models. After analyzing the data empirically, the stages were reduced to only three as abovementioned. The stage called responses and deliberation was integrated with the other three main stages because it was observed that participants continuously provided responses that became a cyclical feedback that fed into the three key stages. Thus the stage called 'responses and deliberation' was no longer considered one stage by itself. In this study, I defined effective participation based on two criteria: (1) *quantity* – number of emails and frequency, and (2) *quality* – substance of emails.

Drawing from Krippendorff's (2004) and Weber's (1990) guidelines, the data analysis procedure was as follows:

- 1. Imported text data files to Atlas Ti** — cleaned data is then ready for coding by multiple coders using the deductive coding technique;
- 2. Determined unit of analysis** — I chose the individual message as the unit;
- 3. Determined coding unit** – I chose to code at sentence and paragraph level within each message;
- 4. Determined the core and sub-concepts to code** — I identified several codes in three distinct categories:
 - Broad category of *situational analysis*, focused on demographics such as message, author's name, date and time message received, and the message author's gender, language use, country/continent/region of origin, and organizational name.
 - Specific category of *effective participation*, focused on contributions made to decision-making process. This is sub-coded as either: 1) problem identification, 2) proposal making, or 3) solution.

² The archival dataset was cleaned with the help of Chia-Jung (Ron) Tsui , a graduate student and team member in COTELCO under the project called 'From Pawns to Partners' at School of Information Studies, Syracuse University. We extend our grateful appreciation to him for his assistance.

³ From April 2003 to July 2005, out of the total messages of 5467, we took out 167 messages that qualify the abovementioned criteria for cleaning the datasets. However, for WSIS Geneva, we only took out 13 messages.

- Specific category of *cultural analysis* focused on *context*, which is divided into two forms – *high context* or *low context*. Sub-concepts and sub-sub concepts include communication styles and action-orientation and are organized in hierarchical dimensions (see Figure 3.5).

5. Generated a coding scheme

Rules for streamlining and organizing the coding process meant that the trained coders coded accurately and thoroughly and ensured that the coding was conducted consistently throughout the text, in the same way every time.

6. Coded the text

6.1 Conducted broad *situational analysis* to provide background or demographic information about Civil Society Plenary participants' email participation, including:

- a. Number of messages generated by month and in total ;
- b. Number of messages generated by individual;
- c. Number of messages received from certain countries, i.e. Mauritania vs. Argentina, or from continent/region, i.e. African region vs. Asian region.
- d. Gender (Male vs. Female) of participants.

6.2 Conducted a specific analysis on *effective participation* (decision-making activities) to identify contextual information related to:

- a. Topic of activities, i.e. policy, speakers, technology, structure, etc.;
- b. Number of activities,
- c. Overall frequency of activities,
- d. Substance of the activities.

6.3 For each selected type of activity, conducted a broad analysis of *culture*.

The purpose of this phase was to answer the question “What patterns of cultural variation are evident in the email messages based on the writer’s high-context versus low-context cultural orientation?” In order to answer this, under each category I defined sub-categories that described the characteristics of high and low context at a lower level of granularity (see Appendix 1). This enabled me to do descriptive statistical analysis such as category counts in terms of frequencies of occurrences (see the following chapter). The purpose of this analysis was to illustrate shared patterns of behaviors exhibited by high-context versus low-context participants.

6.4 Conducted an in-depth *cultural analysis* focused around the three decision-making activities:

a. *Problem Identification*

- This concerns messages in which participants identified a problem(s) or issue(s). Some of the issues were in a form of question, while some were in a form of statement. This activity is crucial as it sets the initial tone for member participation; a well-stated problem is more easily solved than a poorly identified one.

b. *Proposal making*

- This concerns how people generated ideas, as shown by their presenting an idea or making suggestions to participants. From this behavior, I further looked for shared patterns of behavior among high context and low context cultural orientations. For example, high context people sent messages that were lengthy and ambiguous when they were proposing something whereas low context people sent messages that were terse, succinct and direct to the point.

c. *Solution*

- This concerns the manner in which a solution or decision is reached for each of the proposal made, again from a cultural standpoint. Each solution was considered a decision point. For this analysis I looked only at proposals that had a decision. If the proposal did not have a decision point, I regarded it as an instance without a decision or solution.

7. Conducted intercoder reliability check

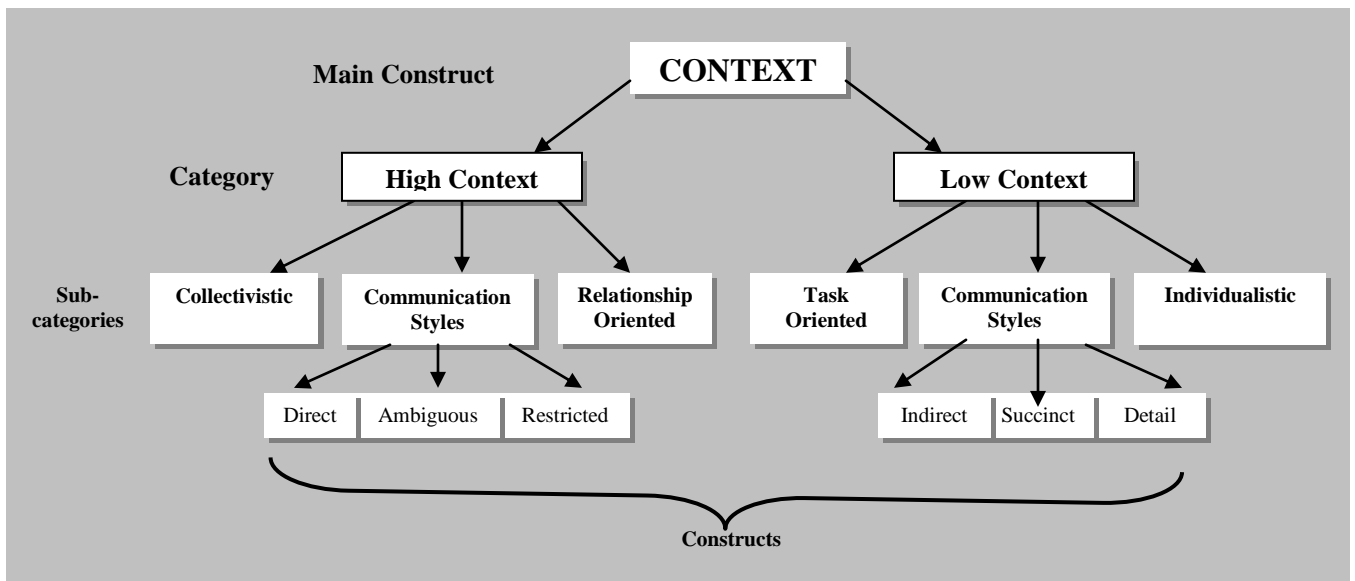
- To increase the reliability of the results, I asked two coders to code the data. The coders were given 13 weeks of training in procedures and use of Atlas Ti including detailed instructions on how to code and what to code in the data. I compared the results of their coding and assessed the level of agreement. My goal was 95% agreement. In the first stage of conducting intercoder reliability test, I achieved a 92% of agreement, but the coders managed to resolve all of their differences.

3.3.4 Data Analytic Framework

This study explored the impact of culture on effective participation of civil society participants during the first phase of WSIS. The data analytic framework was based on deductive coding analysis by using codes generated from theory and literature reviews. In this study, the deductive analysis of the data included (a) identifying the demographic patterns of the individuals, (b) classifying the decision-making process through three decision-making activities (Adler, 1997, Kingdon, 1995) (c) identifying the discernable patterns of intercultural communication styles based on analytic framework of high context and low context (Hall, 1976; Gudykunst et. al., 1996, Triandis, 1994; also see Figure 3.5). Based on the descriptive analysis, a further analysis of culture was conducted which yielded rich in-depth descriptions of the effect of culture on effective participation of Civil Society participants in distributed collaboration during WSIS.

Another important aspect to note is that the cultural analysis relies on basic assumptions derived from high context and low context cultural orientations as well as other related cultural constructs. In other words, the analysis is based purely on the messages that are coded and as such culture is an indicator of the individuals' behaviors as demonstrated in their email messages. In this respect, culture is derived from the messages.

Once the behaviors were coded using the cultural codes, I conducted a second layer analysis by correlating the country variables with the culture variables to tease out whether nationality is a good indicator of high context and low context communication style and cultural value. This analysis was useful to test the validity of Hofstede's cultural framework which assigns context types based on cluster of nationality and not at an individual level. This study on the other hand looked at cultural values at the individual level (e.g. which communication style do people use to make proposal?). This is because with globalization, individuals' behaviors are more context- or situation-dependent and thus not necessarily rooted in their national background. By conducting the second layer analysis, this study explored the assumptions made in studies that use 'nationality' as an indicator of culture.



3.4 Coder Training

In this study, I recruited four coders initially but, in the final stage of coding the whole data set, only two coders were employed because the other coders were no longer available for the task. All the coders were selected based on certain characteristics. As suggested by Krippendorff (1980), "observer, coders, and judges should, of course, be familiar with the nature of the material to be recorded but also capable of handling the categories and terms of the data language reliably" (p. 72). The coders chosen fulfilled this criterion because they were part of the research team in which this study was embedded. As mentioned, this study is part of a larger research lab called 'Collaboratory Technology Enhanced Learning Communities (Cotelco)'. Cotelco had eight members at the time this study was first initiated. The coders

thus were not only familiar with the research context but were also motivated and volunteered to participate in this study because they were personally interested in the phenomenon.

Krippendorff also suggested that coders should have a basic research background in order to increase their competence level. This criterion was met because the coders were graduate students with appropriate training in research or scientific method, though they did not have previous experience in qualitative analysis.

3.4.1 Designing Coding Procedures

Several measurements were taken when designing the coding procedures. The first step was to develop a tentative set of a priori dimensions. These dimensions were based on several different coding schemes selected from the literature reviews (Adler, 1997; Gudykunst, 1996; Hofstede, 1980; Kingdon, 1995; Trompenaars, 1994; Triandis, 1994) and context theory (Hall, 1976). Three sets of different coding schemes were developed in the first stage: (1) demographic codes, (2) decision-making codes, and (3) cultural codes. The tentative dimensions were discussed with colleagues in the research team and some professional colleagues, and as a result the dimensions were further modified for clarity.

Next, the training sessions were organized in order to train the coders.⁴ Coders received approximately 30 hours of training, given in two to three-hour sessions over 14 weeks from September through December 2005 (see Table 3.1). Training included getting familiar with and competent in using the software as well as the codebook. The first step was to educate the coders by providing an overview of content analysis as a qualitative method as well as an overview of Atlas Ti. A full demonstration of Atlas Ti was carried out in the second week of the training.

In the first few sessions, the coders began using Atlas Ti as they applied the demographic codes. The main goal of the first few weeks was to educate the coders to use the software rather than learning the codes. Once the coding scheme was developed, the coders were instructed to begin with the eleven demographic codes (see Appendix 1) because these are less subjective and thus less cognitive effort is needed.

⁴ Training the coders and coding are two of the most challenging tasks in qualitative analysis. A special acknowledgement is given to Cotelco team members who made the task of coding easier. The primary support and encouragement came from the director, Dr. Derrick L. Cogburn, followed by team members (Benjamin Addom, Drew Bennett, Swati Bhattacharyya, Stars Penn, Chia-Jung (Ron) Tsui, Michel Tinguiri, and Antoinette-Marie Uwamwiza).

Table 3.1 : Qualitative Training Sessions (September-December 2005)

Week	Date	Task Completed	Descriptions
Session 1	Sept 20, 05' 10.00am-12.00pm	<ul style="list-style-type: none"> • Presentation: Introduction to Qualitative Method Design and Atlas Ti (see Appendix 5) 	<ul style="list-style-type: none"> • Presentation on qualitative methods and use of Atlas Ti to the research teams including coders
Session 2	Sept 27, 05' 10.30am-12.30pm	<ul style="list-style-type: none"> • Atlas Ti Package Demonstration 	<ul style="list-style-type: none"> • Demonstration of Atlas Ti to research team • Created Demographic Codebook
Session 3	Oct. 4, 05' 11.00am-2.00pm	<ul style="list-style-type: none"> • Atlas Ti Package Demonstration 	<ul style="list-style-type: none"> • Hands-on training on Atlas Ti to coders • Coders began to use the software • Created Decision Making & Cultural Codebook
Session 4	Oct 11, 05' 10.00am-1.00pm	<ul style="list-style-type: none"> • Atlas Ti Package Demonstration 	<ul style="list-style-type: none"> • Hands-on training on Atlas Ti • Coders began to apply basic demographic codes (name, email address, message) to sample data using Atlas Ti
Session 5	Oct 18, 05' 10.00am-12.00pm	<ul style="list-style-type: none"> • Conducted Pilot Study 	<ul style="list-style-type: none"> • Coders began to apply all 3 sets of coding scheme to a sample of 100 messages • Each coders was given 50 messages
Session 6	Oct 25, 05' 10am-12pm	<ul style="list-style-type: none"> • Conducted Intercoeder Reliability Tests (75%) • Discussions and Questions 	<ul style="list-style-type: none"> • Coders discussed their disagreements (25%) • Suggested modifications and clarified lack of understanding on few cultural codes
Session 7	Nov 1, 05' 10am-12pm	<ul style="list-style-type: none"> • Discussions and Questions 	<ul style="list-style-type: none"> • Coders continued to code sample data on paper in order to refine cultural codes • 2 coders dropped out
Session 8	Nov 8, 05' 10am-12pm	<ul style="list-style-type: none"> • Refined and Revised Coding Scheme • Auto-coded decision making stages 	<ul style="list-style-type: none"> • Coders used Atlas Ti remotely and continued coding using the revised codebook • Recruited new coder
Session 9	Nov 15, 05' 10am-12pm	<ul style="list-style-type: none"> • Auto-coded decision making stages 	<ul style="list-style-type: none"> • Coders continued to work on coding sample data for cultural codes • Principle researcher parsed out data for decision making stages using Auto-coding function
Session 10	Nov 29, 05' 10am-12pm	<ul style="list-style-type: none"> • Manual coding of decision making stages 	<ul style="list-style-type: none"> • Principle researcher coded remaining data without auto-coding • Looked for data that was not picked out by auto-coding process
Session 11	Dec 6, 05 11am-12pm	<ul style="list-style-type: none"> • Conducted Intercoeder Reliability for Decision Making stages <ol style="list-style-type: none"> 1. auto-coding (100%) 2. non-auto-coding (100%) 	<ul style="list-style-type: none"> • Coders were given 2 different sets of data to code for decision-making stages—data that was not picked up by auto coding and data that coded using auto-coding • Two coders began to code entire set of data (3 months each)
Session 12	Dec 13, 05 11am-1pm	<ul style="list-style-type: none"> • Questions and clarifications 	<ul style="list-style-type: none"> • Coders continued coding

		<ul style="list-style-type: none"> • Feedbacks 	
Session 13	December 20, 05 11am-1pm	<ul style="list-style-type: none"> • Questions and clarifications • Feedbacks 	<ul style="list-style-type: none"> • Coders continued coding
Session 14	January 9, 06'	<ul style="list-style-type: none"> • Conducted Intercoder Reliability Tests (92%) • Resolved disagreement, reached 100% agreement • Discussions and Questions 	<ul style="list-style-type: none"> • Finished coding the entire dataset (6 months) • Coders provided inputs to the implications and suggestions for the study

Each of the coders was provided with clear written instructions on how to code the data (see Appendix 1). The coders were instructed to apply the demographic codes to a small sample of email messages (months of April and May). Most of the demographic information was located in the email headers such as date and time message was received, name, email address, and gender. By looking at the content or body of the message, coders could identify the author's native language, for example, English, Spanish or French.

By the third session of training, the coders were fully educated and competent to use the software. At this stage, coders moved on to identifying more challenging demographic data such as country, region, organization name and gender. Country name could sometimes, but not always, be identified based on the email extension. Gender posed another problem: for example, the name 'Michel' or 'Sasha' can be either female or male.

In order to deal with these problems, the coders took several steps. First they coded all the doubtful issues or uncertainties with a code called 'coder-question.' They made comments on questionable items, e.g. 'I am not sure of the gender or I am not sure of the country.' Then, a list of the 'coder-question' codes was printed out to identify the names and messages that could not be assigned a country, region, gender, or organization name.

Since the archival email messages were first downloaded to Mozilla Thunderbird, the coders could look up the email extensions based on the 'received' email headers. The email extension provided information for the country codes, and in turn, coders could identify the continent or region. For each message, coders view its routing address which provides a variety of information, some accurate and some misleading.

- The header suffix can identify the name of country directly (de:Germany or ca:canada). Example: sean@nexus.ie, the extension 'ie' directly tells us that the message comes from Ireland.
- The email extension might not tell directly the location of the mail address but might indicate the location of the SMTP server – for example, it can tell that a sender sent his mail from an SMTP mail server located in France. This is misleading information on country of origin of the sender but it does provide some indication.
- The email is registered in Denmark but the sender sent it from Germany. Again, this is misleading information.
- The sender sent the mail via Yahoo or Webmail thus it is difficult or impossible to tell the country of origin.
- The email address ends with an extension of '.org,' and was routed via .net, thus it is difficult or impossible to tell the country of origin.

Another potential and more reliable source of 'country' and 'gender' information was the web-survey conducted by Cogburn (2005) from November 24, 2003 through January 29, 2004. The sampling frame selected in the web-survey consisted of all registered participants in the WSIS virtual Civil Society plenary listserv. This way, the coders could track all the survey participants, match them with the email participants, and then identify the demographic items that they filled out in order to re-code the missing information on country, region, gender, and organization name. However, after looking at this survey based information, some information of the Civil Society was still missing. The last reliable source for the demographic information such as gender, and country/ nationality thus was obtained from the email

registration list. Since I will focus only on 'active' participation, the demographic information that was taken from the registration list was based on the 57 members that were identified as active contributors to the decision making process. But I only obtained 47 people out of the 57 members who were active in virtual Civil Society plenary listserv.

During the fifth session, the coders began to use the full coding scheme which included the decision making and cultural codes. The codes were pre-tested using a pilot sample of 100 messages. The sample data were coded in order to test the coding scheme and some preliminary results such as descriptive analysis (e.g. bar graphs or pie chart) were generated. By doing this, the deductive coding scheme was further added and refined. The process of refinement continued for two more weeks until the coding rules became clear to the coders. In addition, all the challenges faced by the coders were continuously discussed and rectified in the weekly training sessions. The deductive coding scheme was finalized through several iterative processes, and on the eleventh session, the two coders began working on the entire data set. The coders were instructed to apply all the codes to the data to ensure that the categories were exhaustive (Krippendorff, 2004, Neuendorff, 2004).

3.4.2 Revision of Coding Scheme

After eight sessions of training and coding, the coding procedure and coding scheme was revised and refined (see Appendix 2 and 3). For the demographic codes, some codes were extracted while other was modified to include some new codes. For example, instead of just labeling or recording 'message,' the code was further branched out to 'message-sent' or 'message-reply.' The codes were thus tightened and modified to only those codes that were applicable and useful in terms of getting the information. For the decision making codes, only one change was made which was to drop the code called 'responses to ideas.' Instead the coders were asked to identify the three main stages and then each response made in each of the stages would be coded the same stage. For example, if people are responding to problems, then it would be coded as 'problem;' the same rule applies to proposal and solution.

The cultural codes were also refined after referring to existing literature and the outcome of the pre-testing scheme. The sub-categories were reduced to only two codes: cultural values and communication styles. I also decided to use the word 'intercultural communication style' rather than 'communication style' to emphasize the importance of cultural effects and reiterate that the variable is not purely based on one's communication style, but it is rooted in one's cultural background. The two sub-categories are interdependent variables. According to Chen (1997), Gudykunst and Nishida (1997), one's communication styles is deeply rooted in one's cultural values. Under the dimension of high context, the sub-category called cultural values was branched out to two cultural constructs: collectivistic and relationship-oriented, while for low context, it was branched out to individualistic and task-oriented.

In addition, under the intercultural communication styles sub-category, the three existing cultural constructs were reduced to only two. For high context, I retained direct and ambiguous codes. The code called 'restricted' was dropped for two reasons: first, because it was confusing and similar to the meaning of ambiguous; and second, most of the time it is the defining character of ambiguous behaviors. As for low context, the code 'succinct' was dropped because it seemed that succinct was essentially the same as a 'direct' communication style (see Figure 3.5a).

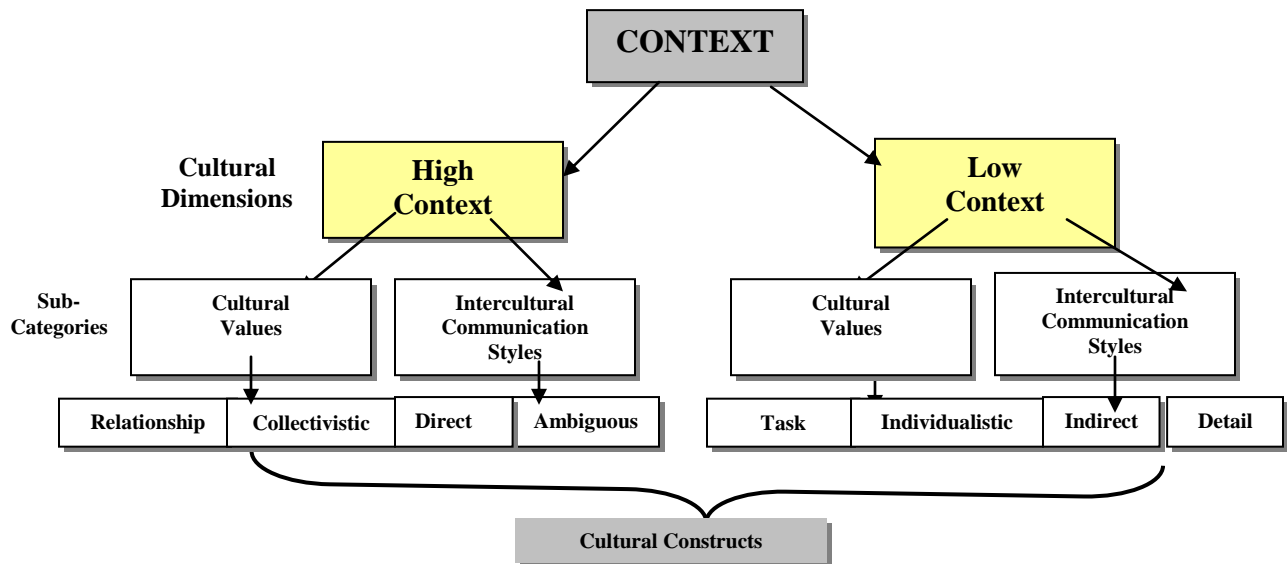


Figure 3.5a. Coding Scheme: Sample of Revised ‘Cultural Context’ codes [adopted from Hall (1976), Gudykunst et al. (1996), and Triandis (1994)]

With the modifications and refinement, the cultural category had eight codes with four codes for each dimension. This resulted in a more parsimonious code scheme but at the same time provided sufficient richness to understand the phenomenon studied. The decision was also made in order to ensure that the categories were mutually exclusive (Neuendorff, 2004).

These revisions were made in order to improve the way the data were to be analyzed. Therefore, several measures (Krippendorff, 2004) were taken to establish reliability in each category:

- The decision to retain or revise the coding scheme was based on the support of existing literature support; codes not supported in the literature were dropped as not worth coding
- The dimensions did not overlap one another
- The dimensions were clear and unambiguous
- The sub-categories under each dimension were mutually exclusive
- The codes were exhaustively applied.

3.4.3 Using Auto-Coding Procedures

Coding the voluminous dataset of email messages (N=1760) was time consuming because most of the messages were long (a range between 2 to 500 words). Since the data needed to be coded from three different aspects – demographic, decision-making, and cultural orientation – coding took even longer. I decided to use the Auto-coding feature in Atlas Ti (see instructions in Appendix 2) as a quicker way to code the data and at the same time speed up my understanding of the nature of the data. Additionally, by using this procedure, it helped me to extract only the decision making activities before the data were given to coders to code using the cultural coding scheme. This procedure is called ‘parsing out data’ (Small, Zakaria & Elfiguigui, 2004).

Based on my immersion in the data for few months prior to the actual coding, I observed that most of the Civil Society participants used words that were indicative of the three decision making stages when they participated in such stages. For example, when identifying problems they often used explicit words like “I have a question about Internet connection...,” or “...the problem is no only that argentina is not a LDC...”. Using auto-coding, I entered the selected word(s) and then applied a code. The auto-code search can be done at six different levels: exact match, word, sentence, single hard return, multi hard return, and all text. I chose to do search at the ‘multi hard return.’ With that function, it captured the word and its entire paragraph. This way, I could read the selected word in its context and it was easier to decide whether to

code or skip each occurrence. The ‘confirm always’ option also allowed me to check the appropriateness of the text before I hit ‘code it.’ Thus, the function operates similar to manual coding but allows the coder to run through the data much faster.

I used auto-coding to capture data for all the three stages of decision making. To increase its effectiveness, I established a list of words that are similar in meaning to the three codes (see Table 3.1a). I referred to a thesaurus and dictionary to come up with an exhaustive list. Using this list, auto-coding was able to pick out almost 65% of the total messages, saving the human coders a great deal of time and effort. The dataset was then given to the two coders who only needed to apply the cultural codes.

Table 3.1a: Words Used by Auto-Coding to Generate Coding Scheme for Decision-Making

Problem	Proposal	Solution
1. Anxiety	1. Application	1. Accept
2. Concern	2. Applications	2. Accepts
3. Concerns	3. Applying	3. Accepting
4. Concerned	4. Arrange	4. Adopt
5. Concerning	5. Arrangement	5. Adopts
6. Conflict	6. Arrangements	2. Adopted
7. Conflicts	7. Idea	3. Adopting
8. Conflicting	8. Ideas	4. Answer
9. Contentious	9. Opinion	5. Answers
10. Issue	10. Opinions	6. Clarify
11. Issues	11. Offer	7. Clarifies
12. Matter	12. Offers	8. Clarified
13. Matters	13. Offered	9. Clarification
14. Problem	14. Offering	10. Conclude
15. Problems	15. Plan	11. Concludes
16. Problematic	16. Plans	12. Concluded
	17. Proposal	13. Consensus
	18. Proposals	14. Decision
	19. Proposal- Making	15. Decisions
	20. Propose	16. Decision-Making
	21. Proposes	17. Explain
	22. Proposed	18. Explains
	23. Proposing	19. Explained
	24. Proposition	20. Explanation
	25. Propositions	21. Explanations
	26. Request	22. Outcome
	27. Requests	23. Outcomes
	28. Scheme	24. Resolution
	29. Schemes	25. Resolutions
	30. Suggest	26. Result
	31. Suggests	27. Results
	32. Suggested	28. Satisfy
	33. Suggestion	29. Satisfies
	34. Suggestions	30. Satisfied
	35. Suggesting	31. Solution
		32. Solutions

Since auto-coding has its limitations, I took careful measures to identify some decision making instances which were harder to code. For example, not all Civil Society participants state problems or express ideas and views using an explicit word or open manner. The words used to describe a problem come in many variations. Some participants used exact matches such as ‘problem,’ ‘issues,’ ‘concerns.’ On the other hand, some problems were phrased in the form of questions. Only through close reading can one

can deduce that it is an expression of a problem in cases such as “Who then do we propose to take any decision on who shall take the civil society speaking slots? Or “who have this list? Is it a secret list? It is like if we do all this job for nothing!”

For the ‘proposal making’ code, similar challenges arose. Many participants used phrases like “I suggest also to add the word “Open” to free software. So, it will be “Open and free software” or “I am suggesting that: 1) The Surveillance Status Report session be documented and translated and shared online to all of us not there...” On the other hand, some proposals were not clearly stated. The most challenging auto-code was that for ‘solution.’ Not only were there few instances of solution activity, but also the types of solutions differed and so people used different terms for the solution-making process. One particular word that Civil Society participants often used was ‘endorses’ or ‘endorsements’ which meant that they agreed with the language of the draft document. The message can be as short as “CFL endorse the document as it is”, or as long and conflicting as “After looking at the drafted documents, we find that we could not endorse the document unless several changes were made in the following paragraph...” Other messages were quite straight forward such as “...the Youth Caucus would definitely support such a proposal.”

After some consideration, I decided to remove the code for “responses and deliberation” stage. Instead I coded that behavior under one of the three distinct stages, again to ensure that the criteria of ‘mutually exclusive’ were met. This made it easier to trace back all the responses and counter-arguments made at each of the stages. There were two ways to identify ‘responses’ received for each of the stages. One way was to read the message and the embedded replied messages. Another way was to look at the message header which gave information such as ‘in-reply-to’ which indicated that the message was sent in response to someone else’s message.

All the abovementioned are examples of messages that would not get picked up by using auto-coding. Therefore, once I auto-coded the data, I further screened the messages that were not found by auto-coding. This gave me another 20% of coded data. Also, to check whether the remaining un-coded data were relevant to any of the decision making codes, I sampled out some messages and asked the coders to code it using the decision making codes. Both coders agreed that the remaining messages were not applicable. By the time the full data set was given to the coders, the revised coding scheme was ready to be used.

3.5 Intercoder Reliability Test

High intercoder reliability guarantees a high level of reliability for the results of a study. Weber (1990) defines reliability as different people being able to code the data in a consistent manner each time they use the same set of instructions. Krippendorff suggests several steps to increase reliability. First, researchers independently review the material and come up with a set of features that form a checklist. Second, the researchers compare notes and reconcile any differences that show up on their initial checklists. Third, the researchers use a consolidated checklist to independently apply coding. Fourth, the researchers compare their coded data and calculate a simple percentage, taking the number of cases coded identically divided by the total number of cases. A 95% agreement is desirable, though some say 75%-80% is still an acceptable level (Krippendorff, 2004). If the level of reliability is not acceptable, then the researchers repeat the previous steps. Once the desired level of reliability has been established, the coding is applied on a large-scale basis. The final stage is a periodic quality control check.

This study conducted two stages of inter-coder reliability testing. Singer (1964) suggested that pre-testing of the coding scheme should undergo intercoder reliability test. Upon arriving at reliability at this stage, the coding scheme can be considered final because the dimensions and categories have been refined and clarified. As mentioned, I had two sets of coders working at two different phases of the data analysis. In the first phase, during the pilot stage, two coders were instructed to apply the preliminary coding scheme to a sample data set of 100 messages. The level of inter-coder reliability agreement was 75%, an acceptable level.

Then, in the following training session, both coders discussed their disagreements regarding coding rules for certain categories. They also discussed the uncertainties and confusions they faced particularly when applying certain cultural codes, such as the difference between ‘succinct’ and ‘direct,’ and between ‘ambiguous’ and ‘restricted.’ The test coders provided some suggestions on how to improve the coding scheme. As a result, I decided to narrow down the subcategories for cultural codes. The coding scheme went through several stages of refinement until the coding scheme became clear; this lengthened

the testing phase but saved time and effort in the long run by establishing a good quality coding scheme from the start.

In the second phase, after the coding scheme was revised, two coders were used; one of them was newly trained, the other had experience from the testing phase. The two coders assessed the dataset according to the three different aspects, demographic, decision-making, and cultural orientation. Each coder independently coded the data. Subsequently they assessed their rate of agreements. With several training sessions, the intercoder reliability increased to 92%. With this percentage, then the coders discussed the differences, and as a result they resolved all confusions and misunderstanding about the codes that they applied (Small, Zakaria & El-Figuigui, 2004). I also asked the coders to validate the auto-coding process by sampling out two sets of data and requested them to code it using the decision making codes. The first set of data contained messages that were parsed out as irrelevant to the decision making process, and the second set of data contained messages that were in relation to the decision making process. The two coders reached total agreement on these two sets of data by using the decision making coding scheme.

Once this level of reliability was reached, the coders used the remaining training sessions for any questions or clarification, and had a chance to discuss problems or suggestions with each other. They also developed a few new rules during these few sessions particularly on the cultural coding scheme. For example, both coders agreed that the 'ambiguity' code was applicable in situations where messages digressed into matters that deviated from the main topic under discussion. They also agreed that 'relationship-oriented' code could be used when people posted extra information regarding topics under discussion, because this activity showed certain aspects of caring and sharing. By the ninth week, the coding scheme was finalized and the coders began using it to code the entire data. Each coder coded three months worth of data and they completed the task in six weeks.

3.6 Chapter Summary

Chapter Three offered a detailed description of the methodology in terms of the design of the empirical study, implementation of the pilot study, testing of the coding scheme, and the data analytic framework. The research setting was the WSIS email listserv from which I collected archival email messages (N=1760); I then carried out content analysis on these messages using Atlas Ti. I used a deductive coding scheme generated from Hall's theoretical lens.

In summary, this study employed an unobtrusive method – content analysis research design – to understand the phenomenon of globally distributed collaboration. The goal of this study was to explore the impact of culture on decision making process when Civil Society members contributed their ideas, views, and opinions using CMC (specifically email). To do this, this study established a deductive coding scheme derived from cultural theories and empirical work and applied it to a massive set of email messages. The next chapter presents the findings this methodology yielded.