

2 **Complementary Explorative Multilevel Data Analysis – CEMDA: A socio-cognitive model of data analysis for Internet research**

Giuseppe RIVA, Carlo GALIMBERTI

Abstract. The Internet can be considered a strategic research site in which to study fundamental social and psychological processes. In fact, online communication systems structure interaction in new ways with dramatic effect on the types of social organizations that emerge from people using them.

Many methods that have been traditionally, and successfully, used in communication research are now being used to study the Internet. Other methods come from anthropology, cultural studies, linguistic, psychology and sociology. How is it possible to choose from and/or integrate all these different approaches?

Starting from a general three-level model of interpersonal interaction in the Web, this chapter tries to define a model of data analysis (Complementary Explorative Multilevel Data Analysis - CEMDA) suited to the constraints of Internet research. Main characteristics of the model are: its roots in the distributed cognition perspective; the focus on different frames and objects for each considered unit of research; the mixed use of quantitative and qualitative tools; the final integration of results in a general framework.

Contents

2.1	Introduction.....	20
2.2	Units of analysis for Internet research.....	21
2.3	Complementary Explorative Multilevel Data Analysis.....	23
2.4	Applying CEMDA to Internet research: choosing and mixing the methods.....	25
2.5	Applying CEMDA to Internet research: the linking of results.....	32
2.6	Conclusion.....	33
2.7	Acknowledgment.....	34
2.8	References.....	34

2.1 Introduction

The Internet can be considered a strategic research site in which to study fundamental social and psychological processes. In fact, online communication systems structure interaction in new ways with dramatic effect on the types of social organizations that emerge from people using them [1]. As recently noted by Jones [2], “The Internet is not only a technology but an engine of social change, one that has modified work habits, educations, social relations generally, and, maybe most important, our hopes and dreams. ... The Internet is a social space, a milieu, made up of and made possible by communication” (p. 2).

As underlined by Mantovani and Spagnoli [3], this can be seen as a general effect of the introduction of new artifacts that disrupt preexisting routines and interactions. In fact, new computer artifacts alter not only the social fabric of communities, but also the kind of relations that tools once had with human minds. These authors also suggest [3] that making sense of new communication environments means “making them part of the sociocultural network that maintains communities and reconfiguring in imaginative ways the existing sociocultural networks” (p. 217).

To describe this new sociocultural network that is the result of communication in the electronic environment researchers and writers are using the term “cyberspace” [4, 5]. In this sense, a key goal for psychology is doing more thinking and theorizing about how to get people to make better connections between the cyberspace and the rest of their lives [6]. However, studying the Internet is not a simple task.

First, the Internet is a medium which can be experienced in many different ways [7]. Though a computer and keyboard are usually the mediator of our Internet experience, there are different ways in which the users can explore the Internet, present themselves and communicate using it.

Second, the Internet is a social and cognitive space [5]. The handling of information is linked to the activation of psychosocial relationships in which cognitions are elaborated. As recently noted by Garton and colleagues [8], “when a computer network connects people or organizations, it is a social network. Just as a computer network is a set of machines connected by a set of cables, a social network is a set of people (or organizations or other social entities) connected by a set of social relations, such as friendship, co-working, or information exchange” (p. 75). Within this network we create relationships but also elaborate cognition. As underlined by Trognon [9], “far from being circumscribed by experimental laboratory settings, cognitive activities are routine daily activities. In other words, it is in everyday life, and in conversational interaction especially, that we put our cognitive skills to practical use” (p.117).

Third, the Internet experience is always situated in a specific context, even when we are chatting alone in a room. In this sense interaction can only be fully understood through detailed analysis of the social context in which it happens [10]: “... at this point we should no longer see people simply as 'users' of given systems, but as social 'actors'. In other words, whether expert computer users or not, people act independently and have their own reasons for what they do, and it is computers and systems that have to adapt to people, not vice versa” (p. 63).

These points clearly underline the situated nature of the Internet experience. More in detail, we can define it as a process by which a group of social actors in a given situation negotiate the meaning of the various situations which arise between them within an electronic environment.

How can we study this complex process? Many methods that have been traditionally, and successfully, used in communication research are now being used to study the Internet. Other methods come from anthropology, cultural studies, linguistic, psychology and sociology.

For instance, McKenna & Bargh [11] suggest that research on the social psychology of the Internet should be organized around three different time phases (before, during, and after extensive social interactions and group participation) and two distinct types of motivations that drive Internet social behavior (self-related and socially related).

According to Tuten & Bosnjak [12] the relationship between personality and Web usage can be studied Using the Five-Factor Model of Personality and Need for Cognition.

Using a sociolinguistic approach, in which social network relations are recognized as the principal vehicle of language change, Paolillo [13] analyzed logfiles of Internet Relay Chat interaction to investigate if and how online interaction have a long-term effect on the evolution of language.

How is it possible to choose from and/or integrate all these different approaches? This is the main goal of this paper. Starting from a general three-level model of interpersonal interaction in the Web, the paper will try to define a complementary model of data analysis suited to the Internet research. Main characteristics of the model are:

- A different focus (general frame and objects included in it) for each of the level considered;
- The mixed use of quantitative and qualitative tools;
- The final integration of results in a general framework.

2.2 Units of analysis for Internet research

The first step in our quest for a new method is to more precisely define what levels/features of the negotiation process need an examination. It's important to note that the definition of the units of analysis doesn't imply that only these units can be profitably used to define a research setting, in isolation of other factors. Indeed, as I noted above, an Internet research study will involve a wide range of considerations, such as social setting and context, user characteristics, or the purpose of the negotiation. But by defining the units of analysis, a researcher can clearly identify what types of processes and settings are under study.

In the previous paragraph, we considered Internet users as social actors with their own aims and autonomy in situations. Particularly the Internet users adapt themselves to the specific situation to achieve their aims. In this sense, social context may regarded as the symbolic system of a given culture which is continually being altered by practical human intervention. Thus:

- *context is conceptual as well as physical*: actors perceive situations using cultural models, and act accordingly in cultural ways;
- *context is unstable*: cultural models are constantly changed by subjects' actions and choices.
- *social context is a prerequisite of communication*: "a shared symbolic order in which action becomes meaningful, and so generates meaning" [10], p.106.

This idea poses serious problems, however [14]. If social actors actively respond to their environment and end up changing it, how can context ever be analysed properly? Mantovani [10, 15] meets the difficulty with a three-level (domain) model of social context which links *the situation* and *social norms* to the *use of computer technology*. The first level is social context in general, the second, ordinary situations of everyday life, and the third, local interaction with the artifact.

The links among the three levels can be studied in either direction, starting from use of computers or from social context. Thus, the use of computers may be regarded as part of everyday life, which is in turn part of the broader social context. By interacting with each

other, the physical environment and the social context, subjects activate a *spiral* of actor-environment exchanges. The First-level person-computer interaction leads to interaction in everyday situations, and thence to cultural changes.

Working in the opposite direction, social context supplies the elements needed to interpret situations correctly, and situations generate the aims which determine a local interaction with the other actors via computer.

Mantovani's model doesn't directly address the mediated communication between different actors. However, it can be easily adapted to arrange all the aspects of the Internet experience. We propose here a revised three-level model - composed by *context*, *situation* and *interaction* - that mutually define the social context in which the Internet experience is situated (see Figure 2.1). These three levels can be considered the basic units of analysis for Internet research.

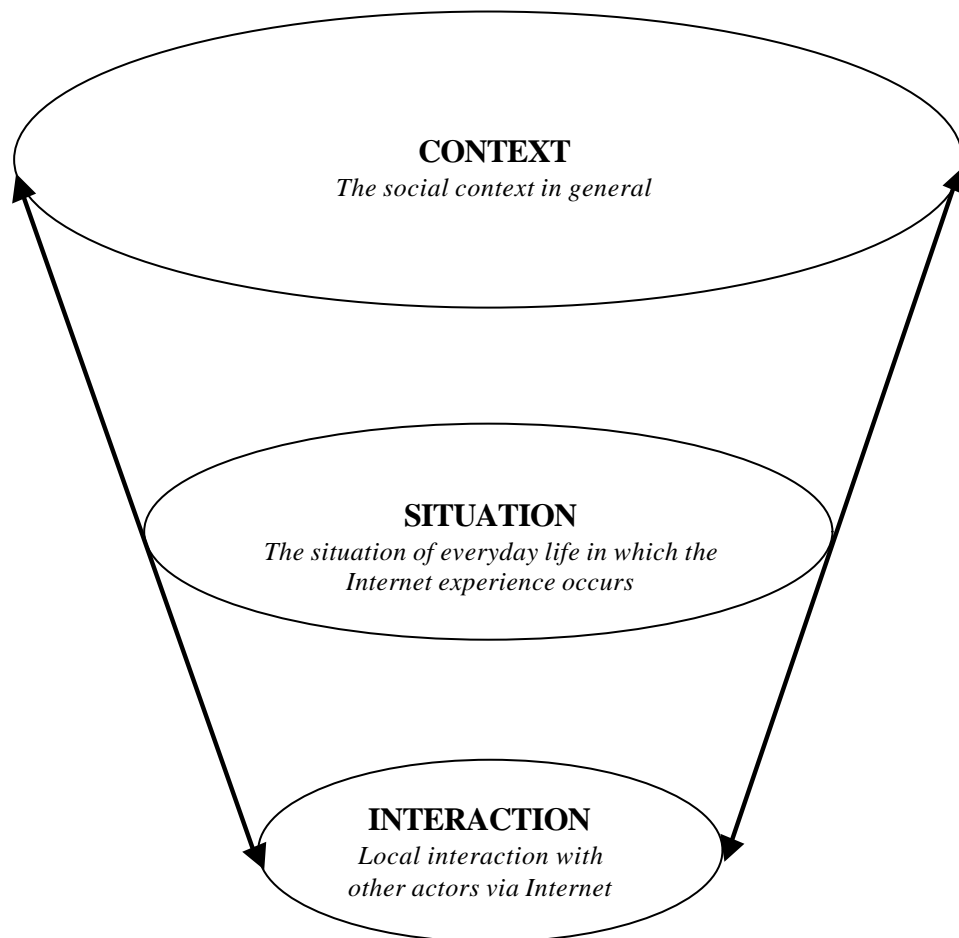


Figure 2.1 A three-level model of the Internet experience

The first level remains social context in general (the context); the second, is the situation of everyday life (the situation) in which the Internet experience happens; and the third, the local interaction with the other actors via Internet (the interaction). This interaction is direct during computer mediated communication (actors are directly experienced), and indirect in other Internet experiences such as browsing a Web site. However, also during Internet surfing, the artifacts experienced reflect the characteristics and goals of the actors who created them.

2.3 Complementary Explorative Multilevel Data Analysis

Identifying the units of analysis is just the first step in our quest for a method. The next one is to define how to study them. At this point, when the methodology needs to be chosen, the qualitative versus quantitative debate begins. As noted by Sudweeks and Simoff [16], “Quantitative and qualitative methods are quite distinct in the emphasis they place on each.

In quantitative analyses, argumentation is based on a representation of the phenomenon as a finite set of variables. There, we seek systematic statistical or other functional relations between these variables. In qualitative analyses, argumentation is based on a description of the research objects or observation units instead of on approximation of a limited number of variables. In other words, in qualitative analyses, references to excerpts or cases in the data are used as clues.” (p. 39).

To overcome this opposition many attempts at integrating the two methods appeared over the past two decades. The final aim is to define a combination of research methods so that the weakness of any single method - qualitative or quantitative - is balanced by the strengths of other methods. In reality, however, the qualitative and quantitative analyses are usually distinct, mutually exclusive components of the research. A possible sample of this approach is reported by Sudweeks and Simoff [16]: “One component is unstructured textual data of a phenomenon being investigated (e.g., transcripts of interviews or verbal reports from protocol studies), analyzed with an interpretative or hermeneutic method. The other component is numerical data of the same phenomenon (e.g., from a content analysis or a survey questionnaire), analyzed with some statistical procedure. The result is an integrated view that narrowly focuses on a particular social phenomenon” (p. 40).

A further improvement over this integrated approach is the appearance of the Complementary Explorative Data Analysis (CEDA) framework [16]. CEDA incorporates complementary use of both methods, depending on the particular research stage or the initial assumptions that need to be taken into consideration. More in detail, the first stage is devoted to the identification of domain specifics, the applicable research methods and the possible research aims. Once the scope is specified, CEDA follows the following processes: after the data mining and the qualitative reasoning, qualitative and quantitative methods are used to collect data; the results obtained are tuned using a qualitative refinement.

Note that the data collected in any of the research domains are a combination of quantitative measurements and qualitative observations. Specifically, CEDA uses quantitative methods to extract reliable patterns, whereas qualitative methods are incorporated to ensure capturing of the essence of phenomena. In this approach the use of different data sets in the same research is allowed.

In this paper we propose a Complementary Explorative Multilevel Data Analysis (CEMDA) framework that applies the CEDA approach to the three-level model used for the analysis of the Internet experience. CEMDA has the potential to conduct parallel and interconnected research on the same domain. In particular, CEMDA allows the researcher to focus on the different levels of the Internet experience. As we have seen before, the proposed model for the analysis of the Internet experience is structured on three different levels:

- *the context*: social context in general;
- *the situation*: ordinary situations of everyday life in which the Internet experience occurs;
- *the interaction*: local interaction with the other actors using a technological artifact (Internet).

How can CEMDA approach each of them? The first step is the definition of domain specifics. Particularly, for each of the three levels the researcher has to identify the general frame of analysis and the objects included in it (see Figure 2.2). Let's use an example to explain this point.

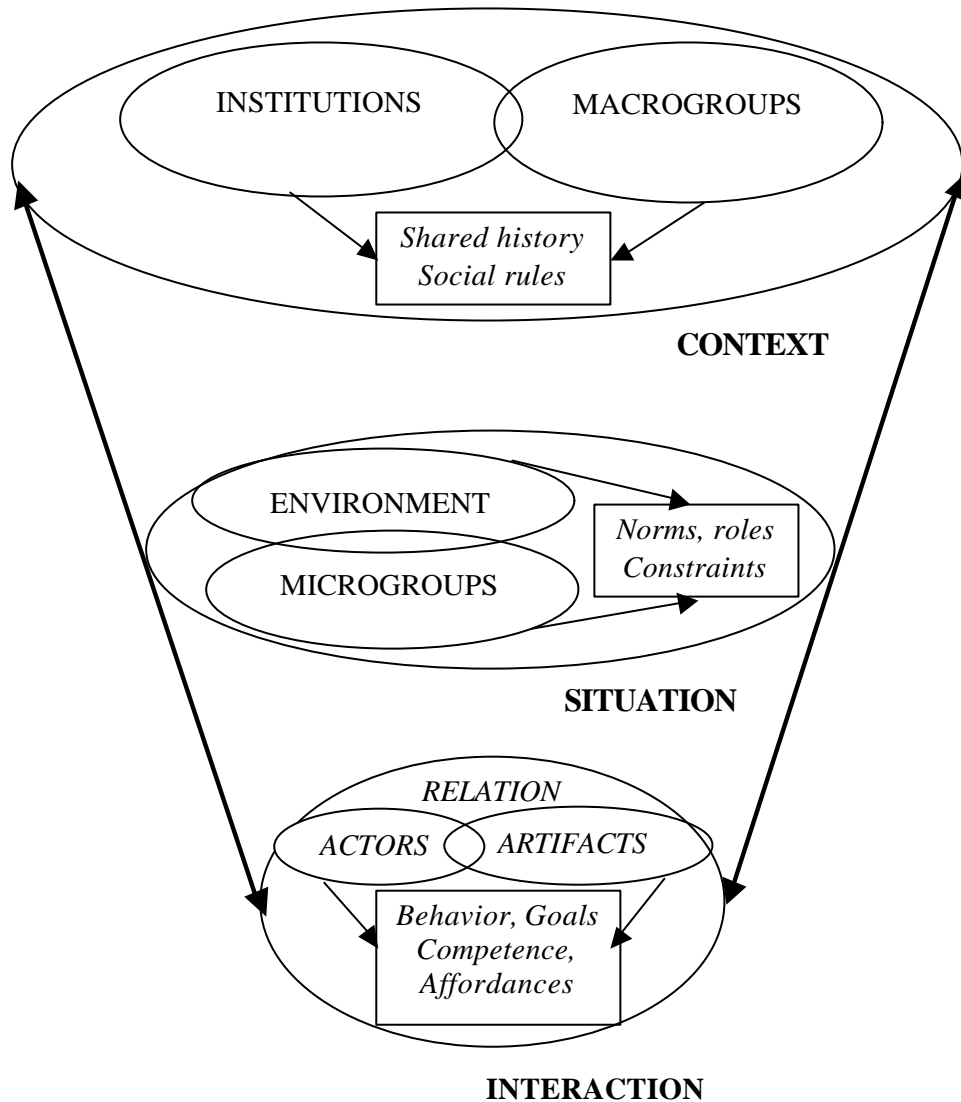


Figure 2.2 Frames and objects in the three-level model

A researcher is interested in studying how the doctors in different hospitals use e-mail for preparing and sharing clinical diagnoses. The starting point of the analysis will be the identification of the different frames and objects of analysis. If we start from the context, the frames analyzed are both the institutions that shape the social context - in this case the hospital - and the macrogroups which interact in them: the different departments/institutes.

Specifically, the objects included are the history of the hospitals - their backgrounds, how their collaboration started, etc. - and the social rules structuring the general interaction processes: how departments decide, when, how it is possible to obtain resources, etc.

The second level of analysis is the situation in which the Internet experience is located. In the example presented the frame includes all the doctors (microgroup) directly involved in the definition of the clinical diagnosis and the social/physical environment in which the process happens. Objects analyzed by the researcher are the roles within the microgroup -

how is structured the diagnosis process, the roles of the different actors, etc. - and more in particular the norms that regulate the interpersonal interaction in it: who starts the diagnosis process, who takes the final decision, etc.

Also, relevant objects are the environmental constraints (social and physical) that can limit the Internet experience. We can find communities of practices or the availability of the tools (e.g., I can use them only when the head of the department is present), on the social side. Bandwidth or a limited connection time is typical constraints on the physical side.

The final level of analysis is the interaction. The general frame of the analysis is the relation process. Specifically, how the actors interact between them using the artifacts. The first object explored by the researchers is the reciprocal goals of the actors - integrating in a coherent diagnosis the information received, doing it in the less possible time, etc. After understanding them, the analysis moves on the affordances perceived in the situation and the availability of the competencies needed to exploit them: e.g., opening the image it is possible to identify possible disturbances; but is the doctor able to open the images received? Final objects are all the behaviors, including the communicative ones, produced by the actors in their interaction with the artifacts: switching on the computer, opening the e-mail software, the production of text, etc.

2.4 Applying CEMDA to Internet research: choosing and mixing the methods

2.4.1 The choice of the overall approach to the research

For each of the frames and objects identified, the researcher has to identify the more suitable methods for the analysis [16]. What is the starting point?

The first step is the choice of the overall approach to the research. Most researches are based on two approaches: studies aimed at discovering causal relationship, and descriptive studies that use collected data to describe a phenomenon [17].

Table 2.1 The possible approaches to the research

<p>Experimental approach</p> <ul style="list-style-type: none"> - can truly test hypotheses concerning cause-and-effect relationships; - requires <i>randomized</i> groups (not required in quasi-experiments); - requires the control of many variables and the possibility of manipulating at least one of them (independent variable/s); - requires the control of <i>internal validity</i> (the changes observed in the dependent variable are due to the effect of the independent variable) and <i>external validity</i> (the extent to which findings in one study can be applied to a different context) by means of <i>research designs</i> (i.e. pretest-posttest control group design). - requires a specific statistic method: t-test, ANOVA, MANOVA, non-parametric tests. <p>Causal Comparative approach</p> <ul style="list-style-type: none"> - the results cannot be used as a proof of a cause-and-effect relationship; - requires at least two different groups (defined group and comparison group); - requires a clear basis for categorization (i.e. gender, or age); - requires a specific statistic method: t-test, ANOVA, MANOVA, non-parametric tests. <p>Correlational approach</p> <ul style="list-style-type: none"> - the results cannot be used as a proof of a cause-and-effect relationship; - several variables can be easily included in one study; - requires the identification of <i>exploratory/predictor</i> variables and <i>outcome/criterion</i> variables; - requires a specific statistic method: simple correlation, regression analysis, multiple regression analysis, discriminant function analysis, canonical correlation, path analysis, and factor analysis.
--

In general causal studies require experimental (subjects randomly assigned to the groups) or quasi-experimental (subjects are not randomly assigned to the groups) designs in which the researcher can manipulate the independent variable/s (see Table 2.1). However, given the peculiar characteristics of the Internet, in many situations this is impossible for the researcher. In these settings the researcher can use a causal comparative or a correlational approach (see Table 2.1).

There is a significant difference between these two approaches. If correlational approach is used when the goal of the researcher is to investigate the magnitude of the relationship between two or more variables (i.e. the dimension of a banner and the click-through rate), the causal comparative approach allows the comparison of individuals/groups who differ on a specific variable (i.e. how genders differ in the click-through rate). However, given their characteristics neither the correlational nor the causal comparative results can be used as a proof of a cause-and-effect relationship.

2.4.2 *The choice between quantitative and qualitative methods*

The second step is the choice between quantitative and qualitative methods: the researcher can use just one or both of them. The general guideline for this choice is the final goal of the analysis. Generally, qualitative methods are used to understand an observed phenomenon, while quantitative methods aim at explaining it.

Another important point to discriminate between the two approaches is the starting point of the analysis. Qualitative analysis usually begins with a theory formulated as a set of hypotheses, and the purpose of a study is to find support for or to disprove the theory. Quantitative research is instead used to provide the ability to predict, and control examined concepts. To reach this goal, the researcher needs to start from the design of an experiment. This calls for the knowledge of the form, type, and range of the content of the data before the start of the analysis.

A third point to identify the right method is the role played by the researcher. In qualitative analysis the role of the researcher is participatory and personal. This is usually needed when the final goal of the study is to improve or tune a specific Internet experience - i.e. improve the characteristics of a Web site. On the opposite, quantitative methods require that the researcher acts just as an observer with a limited control of the setting. If in a passive experiment, the researcher just records the observations, in an active experiment the researcher may need to set up and control some of the independent variables.

As noted by Sudweeks and Simoff [16], the issue which these approaches differ most is the priority placed on the role of interpretation. In every research some form of interpretation is required. But, whereas quantitative research calls for the suspension of interpretation during the experiment, qualitative research actively interprets phenomena during the observation phase.

When does the researcher mix quantitative and qualitative methods? Mixed methods are usually found in two situations:

- lack of a clear starting point for the analysis: no theory to confirm or experiment to design. In this situation, typical of explorative analysis, the merging of methods can offer to the researcher a broader data set;
- multi-step analyses: after testing a hypothesis using qualitative methods, the researcher tries to control examined objects in an experimental settings. This approach is usually used when the final goal of the researcher is to modify/improve the original Internet experience.

After identifying the right approach, the next step is the definition of the methods to be used. In fact, each approach has to be targeted to the level considered. In Table 2.2 specific

qualitative and quantitative methods are reported for the different levels: context, situation and interaction. A more detailed description of the methods reported can be found in the next paragraphs.

Table 2.2 Units of analysis and methods used within the Complementary Explorative Multilevel Data Analysis (CEMDA) Framework

Levels	Frames	Objects	Quantitative analysis	Qualitative analysis
<i>Context</i>	Institutions, Macrogroups	Shared history, social rules	Survey questionnaires	Social network analysis, Interviews, Document and records review
<i>Situation</i>	Microgroups, Social and physical environment	Norms, Roles, Environmental constraints, Practices	Survey questionnaires	Social network analysis, Interviews, Focus groups, Document and records review
<i>Interaction</i>	Relations, Actors, Artifacts	Goals, Competence, Affordances, Behaviors (including the communicative ones)	Individual questionnaires, Quantitative interaction analysis (duration and structure, computer monitoring, eye tracking, etc.)	Qualitative interaction analysis (Speech acts analysis, discourse analysis, analysis of conversations, etc.) Participant observation, Interviews, Diaries

2.4.2.1 Survey and questionnaires

A typical quantitative tool used in CEMDA is the questionnaire. Usually the researcher has three different possibilities [17]:

- to use a measurement instrument that is commercially available or that has been developed by other researchers;
- to adapt an existing instrument;
- create a new one to meet the needs of the proposed research.

Given the limited number of questionnaires targeted to the Internet experience, it is possible that the researcher will be forced to develop his/her own data collection instrument. The steps outlined here give a rough guide to this complex process [18].

The first step is the *definition the objective of the instrument*. To reach this goal the researcher has to answer to these general questions: what is the specific purpose of the proposed measurement instrument? What information do you want on what attribute?

After the identification of the overall goal of the instrument the next step is the *review of existing measures*. The previous work conducted by other researchers can help in finding methods for formatting and administering the measure as well as for determining its reliability and validity.

The third step is the *identification of the possible respondents*. The choice of the possible respondents allows the definition of key factors related to the administration of the instrument: the amount of time required to complete it, reading level, format for items, response option formats, and test setting. Format options usually include true-false, matching, multiple-choice, sentence completion, ranking items, Likert-type scales (1:

strongly agree; 2: moderately agree; 3: neutral; 4: moderately disagree; 5: strongly disagree), and open-ended, essay-type questions [17].

The fourth step is the *definition of an appropriate title for the questionnaire*. This is the first thing that the respondent will read, so it has to be motivating and conceptually consistent with the content of the instrument. Following the title, is usually included a short, introductory paragraph explaining the purpose of the instrument and its intended use.

The fifth step is the *definition of pool of possible items*. Usually some items are adopted or adapted from similar questionnaires. Others might be developed using experts or program staff responsible for the program being studied.

After the identification of the items the next step is the *pilot test of the questionnaire*. To develop a good instrument, different pilot tests are required. At first, it is recommended that the developer ask other professionals expert in the field to review the prototype. After a first revision, the prototype can be tried out on a small sample of the intended respondents. Typically, this is done by the researcher under expected administration procedures to get a general idea of the quality of the information as well as any problems in administration and scoring. The final pilot test should be conducted with a larger sample. The goal of this test is to enable the researcher to gather reliability and validity information.

The last step is the *final revision of the instrument using an item analysis*. The answers to each item should be reviewed to verify if a pattern suggests ambiguity or bias in the item.

2.4.2.2 Observation

Observation is a qualitative method occurring in naturalistic settings without the use of predetermined categories of measurement or response [19]. Adler and Adler [19] distinguish between observation and participant observation: in participant observation the researcher usually interacts with the participants while collecting data from them. According to Spradley [20] it is possible to identify five levels of participation:

- *Non-participation*: The lowest level of involvement is usually obtained by watching a through a unidirectional mirror or by examining a videotape.
- *Passive participation*: The researcher is present but does not interact with the participants.
- *Moderate participation*: The researcher attempts to balance the insider and outsider roles by observing and by participating in some but not all of the activities.
- *Active participation*: The researcher does what the others do, generally, but does not try to blend in completely.
- *Complete participation*: The researcher becomes a natural participant, which has the disadvantage of trying to collect data and maintain a questioning and reflective stance.

What has to be observed during a session? According to Patton [21], key features are:

- *the physical environment*: the amount of space available, how the space is used, the nature of the lighting, how people are organized in the space, and the interpretive reactions of participants to the physical setting.
- *the social environment*: the focus is the way in which participants organize themselves into groups and subgroups. In particular have to be examined the patterns of interaction, frequency of interaction, direction of communication and decision-making patterns.
- *nonverbal communication*: dress, expression of affection, physical spacing, and arrangements are typical nonverbal cues. Other important cues are the patterns established for the participants to get the attention of or approach another person.

Finally, also fidgeting, moving about, or trying to get comfortable can provide information about attention to and concentration on group processes.

- *observing what does not happen*: If the session has a specific goal suggesting that certain things are expected to happen, it is appropriate for the observer to note what things did not happen.

2.4.2.3 Interview

As we have just seen, observation allows collection of data through the researcher's direct contact with the setting of the research. However, it is not always possible to have intimate, repeated, and prolonged involvement in the setting of the research [22]. To solve this problem, usually researchers plan interviews with key subjects.

Interviews can be structured or unstructured, group or individual. Typically, interviews in a qualitative study are done with an unstructured or minimally structured format. Interviewing can be conducted as a part of participant observation or even as a casual conversation. As the study evolves, interviewing can become more structured and formal.

It is also possible to use questionnaires in interviewing. As noted by McCracken [22] "The questionnaire has several functions. Its first responsibility is to ensure that the investigator covers all the terrain in the same order for each respondent. The second function is the care and scheduling of the prompts necessary to manufacture distance The third function of the questionnaire is that it establishes channels for the direction and scope of discourse. The fourth function of the questionnaire is that it allows the investigator to give all his or her attention to the informant's testimony" (pp. 24-25).

2.4.2.4 Focus group

Focus groups, can be seen as group interviews that rely, not on a question-and-answer format of interview but on the interaction within the group [17]. The key role of the interaction between participants is designed to elicit more of the participants' points of view. In fact, the main advantage of focus group research is the additional insight gained from the interaction of ideas among the group participants.

The use of focus groups is appropriate when the researcher is interested in how individuals form a schema or perspective of a problem. In this sense the focus group can be considered a "guided" discussion, in which the facilitator usually has a list of five to seven questions to ask during a 1.5- to 2-hour session. The questions are used in a semi-structured way to ensure coverage of important issues allowing for flexibility in responding to group-initiated concerns.

A key variable for the researcher is the number of groups to be included in the study. According to Morgan [23] only three or four groups are necessary when the research is highly structured and exploratory. However, he recommends the use of six to eight groups if the goal is a detailed content analysis with relatively unstructured groups.

2.4.2.5 On-line focus group

Usually focus groups occur in a face-to-face setting. However, the latest and most interesting forms of synchronous computer mediated communication – *shared hypermedia* – now allow the set up of web-based focus groups.

Shared Hypermedia are new Internet tools attaching computer mediated communication to Web browsing (see Table 2.3 for a list of available shared hypermedia). In shared hypermedia, different users, who are simultaneously browsing the same Web site, can communicate with each other and share files or web addresses.

Table 2.3 Shared hypermedia tools

Tool	Developer	Website
<i>Cahoots</i>	Cahoots	http://www.cahoots.com
<i>Instant Rendezvous</i>	Multimate	http://www.multimate.net
<i>Odigo</i>	NovaWiz	http://www.odigo.com
<i>ICQSurf</i>	ICQ	http://www.icq.com/icqsurf

Using a simple interface, usually resembling a little remote control, shared hypermedia users can get a constantly updated list of all the other online users who are visiting the same Web site. All they have to do is click on any person icon, open a message window, and start the communication.

Usually a shared hypermedia lets the user conduct group and private chats, exchange information or files, and share the same web pages (for a more detailed characteristics of a typical shared hypermedia see Table 2.4).

Table 2.4 Features of shared hypermedia tools

One-on-one or multi-user chat (text, audio and video)	<i>Users can make calls to multiple people up to 100/1000 users. In multi-user chat one or more moderators can control group participation by sharing the microphone. It is also possible to broadcast a radio-style Internet talk show where the host maintains control and invites listeners to participate to the event.</i>
Email (text and voice)	<i>User can send text and voice messages to users who are not online.</i>
Web tour	<i>Users can create their own Web Tour and escort other users through a list favorite web sites.</i>
Search engine	<i>Users can find other users with a specific sex, age and/or similar interests.</i>
Transfer of files	<i>Users can upload and download from other users documents and files.</i>

On any Web site, shared hypermedia users can see a list of other users and talk with them on group and private levels. Shared hypermedia further enhances the user experience by consolidating different form of computer mediated communication (e-mail, Internet Relay Chat, etc.) into one fully integrated interface. Many shared hypermedia also have a search engine that can be used to find users with a specific age and/or similar interests. In this way it is really easy to set up a group with a common interest, like medical oriented web sites.

The most advanced shared hypermedia (i.e., Cahoots) have an option - the *web tour* - very interesting for its possible use in Internet oriented focus groups. During a web tour a facilitator can guide the browsing of a small group of users (usually up to 20), who can discuss and interact between them in real time.

2.4.2.6 Documents and record review

Usually every organization has huge amounts of documents and records that describe its history and current status. In general, it is possible to define a *record* any output of the

organization prepared for official reasons (i.e.: bank statements, file records, etc.). A document is any output of the organization prepared for personal reasons (i.e.: letters, memos, etc.). Documents and records include not only the typical paper products, such as memos, reports, and plans, but also computer files, tapes (audio and video), and other artifacts [24].

The Internet researcher must analyse these documents and records to get the necessary background of the general context and insights into the dynamics of everyday functioning.

In particular the researcher has to be sensitive to the types of records and documents that might be associated with a particular context or situation. Must be noted that the use of these materials has to be tempered with an understanding of the time, context, and intended use for which the materials were created. In many situations it may not be possible to interact with the people who produced the materials. The researcher then faces the challenge of how to interpret the meaning of such materials.

2.4.2.7 *Social network analysis*

Social network analysis focuses on the relations among actors, and not individual actors and their attributes [25]. This means that the subjects are usually not sampled independently, as in many other kinds of studies (most typically, questionnaires). In particular, network analysts study the patterning of the social connections that link sets of actors. For the most part they seek to uncover either or both of two kinds of patterns. They often look for *social groups* -- collections of actors who are closely linked to one another. Or, alternatively, they look for *social positions* -- sets of actors who are linked into the total social system in similar ways [26].

So, if the nodes or actors included in non-network studies tend to be the result of independent probability sampling, network studies are much more likely to include all of the actors who occur within some boundary. Suppose we are studying friendship ties developed in a chat, for example. John has been selected to be in our sample. When we ask him, John identifies three friends he knew in the chat. Network analysis requires to track down each of those three friends and ask them about their friendship ties, as well. The three friends are in the sample because John is (and vice-versa), so the "sample elements" are no longer "independent." In this sense, within social network analysis, a group is an empirically discovered structure: it emerges by the pattern of relations among members of a population [8].

Given a set of actors, there are several strategies for deciding how to go about collecting measurements on the relations among them [25]. The broadest but more demanding approaches are "full network" methods. Full network methods require that the researcher collects information about each actor's ties with all other actors. In essence, this approach is taking a census of ties in a population of actors -- rather than a sample.

At the other end of the spectrum there are "ego-centric networks". In this approach the researcher begins with a selection of focal nodes (egos), and identify the nodes to which they are connected. Then, he determines which of the nodes identified in the first stage are connected to one another.

Information about the ties is usually gathered by a combination of methods including questionnaires, interviews, diaries, observations and more recently through computer monitoring. People are also asked to identify the frequency of communication with others as well as the medium of interaction.

To analyze these data researchers usually used specific designed software: *Multinet*, *Negopy*, *Krackplot*, and *Gradap* (see Table 2.5), with the combination of *UCINET* and *Krackplot* being the mostly widely used [8]. To use these applications, data often must be transformed into a matrix with rows and columns representing the units of analysis (people, events, groups or other entities that are related to one another).

Table 2.5 Software for Social Network Analysis

Tool	Developer	Website
<i>Multinet</i>	Vancouver Network Analysis Team	http://www.sfu.ca/~richards/Multinet/Pages/multinet.htm
<i>KrackPlot</i>	Analytic Technologies	http://www.contrib.andrew.cmu.edu/~krack/
<i>Negopy</i>	Vancouver Network Analysis Team	http://www.sfu.ca/~richards/Pages/negopy.htm
<i>Gradap</i>	ProGAMMA	http://www.gamma.rug.nl/files/p315.html
<i>UCINET</i>	Analytic Technologies	http://eclectic.ss.uci.edu/~lin/ucinet.html

2.5 Applying CEMDA to Internet research: the linking of results

This complementary analysis is based on the linking of the results obtained by each of its components. Specifically, the main advantage of CEMDA is the possibility of using the data collected on one of the level to tune/define the characteristics of the analysis in the other levels. How can this be done?

As underlined by Mantovani [10] the links among the three levels can be studied in either direction, starting from the mediated interaction between the actors or from the general social context. Usually, the researcher identifies the starting level according to the general aims and hypotheses of the study. For instance, if the goal of the research is to identify how to improve the diagnosis process using the Internet, the starting point can be the interaction between the actors during the actual diagnostic tasks.

On the opposite, if the researcher wants to explore the effect of the Internet-based diagnostic process on the organization of the hospital, the general social context could be the starting point.

However, in both cases, the main characteristic of CEMDA is to analyze a level without forgetting the possible influences of the remaining ones. In the first example the diagnosis process has to be considered as part of everyday life of the department considered, which is in turn part of the broader social context of the hospital. More in detail, this can be done using the following approach (see Table 2.6 for a detailed breakdown):

- definition of the starting level. For example, the interaction between the actors during the actual diagnostic tasks.
- identification of the links between the starting level and the previous or next one. In particular, *if* and *how* a specific frame in the starting level is connected to the objects of the next one (e.g., the relation between the actors is related to the constraints of the environment).
- when a direct link is identified, the researcher has to understand its *direction* (cause and effect). For instance at the *situation* level, the physical constraints of the environment (e.g., the location of the networked computer) may influence the characteristics of the relation between the actors (e.g., too much time is needed to reach the computer, so the interaction time is limited) in the *interaction* level.
- if a frame is causally influenced by an object found in a different level, the data coming from the other level will be used in the definition of methods (e.g., use of survey and interviews to explore the constraints of the environment) and in the final discussion (e.g., actual interaction time is limited, but this can be explained by the wrong location of the networked computer).
- the final results may lead to revision of the identified domain (e.g. we can decide to investigate why the networked computer is in the actual location) and changes in the combination of analysis methods. Specifically, the results can be used to produce

new hypotheses (e.g., if we move the computer to a better position, the interaction time will be longer) or to define new experiments (e.g., we can measure interaction time using computers in different positions).

Table 2.6 Stages of Complementary Explorative Multilevel Data Analysis

- Scope of research
- Levels identification
- Definition of a starting level
- Definition of frames and objects (starting level)
- Definition of frames and objects (other levels)
- Identifications of links between the starting level and other levels
- Analysis of the links (direction)
- Definition of hypotheses or experiments (starting level and causally connected levels)
- Definition of methods (starting level and causally connected levels)
- Data Collection and Selection of the Data Sets (starting level and causally connected levels)
- Eventual integration of quantitative and qualitative results (starting level and causally connected levels)
- Integration of data from different levels
- Interpretation and presentation
- Eventual production of new hypotheses or formulation of a new experiments

2.6 Conclusions

As we have seen in the Introduction, studying the Internet is not a simple task. First, the Internet is a medium which can be experienced in many different ways. Second, the Internet is a social and cognitive space. Third, the Internet experience is always situated in a specific context, even when we are chatting alone in a room. In this sense it can only be fully understood through detailed analysis of the social context in which it happens. These points clearly underlined the situated nature of the Internet experience, defined as a process by which a group of social actors in a given situation negotiate the meaning of the various situations which arise between them within an electronic environment.

In this paper we tried to address a very simple problem: how can we study this process? The starting point in our quest for a method was the identification of the units of analysis. Using a three-level model of the Internet experience we identified three macro units (levels): the social context (the context); the situation of everyday life (the situation) in which the Internet experience happens; and the local interaction with the other actors via Internet (the interaction).

To explore these units we proposed a new approach, the Complementary Explorative Multilevel Data Analysis (CEMDA) framework, allowing the conduction of parallel and interconnected research on the same domain. This complementary analysis is based on the linking of the results obtained by each of its components.

As with all research methods, however, the CEMDA framework is only a tool for simplifying and understanding rather complex data, and are not a substitute for insight, clear thinking, and intimate knowledge of the subject matter. Properly used this approach will provide researchers interested in the area of with powerful and more efficient analytic tools for examining the characteristics of the Internet experience. Specifically, the main advantage of CEMDA is the possibility of using the data collected on one of the level to tune/define the characteristics of the analysis in the other levels.

2.7 Acknowledgment

A preliminary version of this chapter was originally published with the title “The mind over the Web: The quest for the definition of a method for Internet research” by the journal “*CyberPsychology & Behavior*”, 4 (1), 7-16, 2001. Copyright Mary Ann Liebert, Inc., 2001.

The present work was supported by the Commission of the European Communities (CEC), in particular by the IST programme (Project VEPSY UPDATED, IST-2000-25323 - <http://www.psicologia.net>; <http://www.vepsy.com>).

2.8 References

- [1] P. Kollock and M. A. Smith, Communities in cyberspace, in *Communities in cyberspace*, M. A. Smith and P. Kollock, Eds. London: Routledge, 1999, pp. 3-25.
- [2] S. Jones, Studying the Net: Intricacies and issues, in *Doing Internet research: Critical issues and methods for examining the Net*, S. Jones, Ed. Thousand Oaks, CA: Sage Publications, Inc., 1999, pp. 1-27.
- [3] G. Mantovani and A. Spagnoli, Imagination and culture: What is it like being in the cyberspace?, *Mind, Culture, & Activity* 7 (2000) 217-226.
- [4] L. Strate, R. Jacobson, and S. B. Gibson, Communication and Cyberspace: Social interaction in an electronic environment. Cresskill, NY: Hampton Press, 1996.
- [5] S. Kiesler, Culture of the Internet. Mahwah, NJ, US: Lawrence Erlbaum Associates, Inc., Publishers, 1997, pp. xvi, 463.
- [6] S. Turkle, Parallel lives: Working on identity in virtual space, in *Constructing the self in a mediated world. Inquiries in social construction*, T. R. L. Debra Grodin, Ed.: Sage Publications, Inc, Thousand Oaks, CA, US, 1996, pp. 156-175.
- [7] P. Wallace, *The psychology of the Internet*. New York: Cambridge University Press, 1999.
- [8] L. Garton, C. Haythornthwaite, and B. Wellman, Studying on-line social networks, in *Doing Internet research: Critical issues and methods for examining the Net*, S. Jones, Ed. Thousand Oaks, CA: Sage Publications, Inc., 1999, pp. 75-105.
- [9] A. Trognon, Psicologia cognitiva e analisi delle conversazioni [Cognitive psychology and conversation analysis], in *La conversazione. Prospettive sull'interazione psicosociale [Conversation: perspectives in psycho-social interaction]*, C. Galimberti, Ed. Milan: Guerini e Associati, 1992, pp. 110-122.
- [10] G. Mantovani, *New communication environments: from everyday to virtual*. London: Taylor & Francis, 1996.
- [11] K. McKenna, Y.,A. and J. Bargh, A., Causes and consequences of social interaction on the Internet: A conceptual framework, *Media Psychology* 1 (1999) 249-269.
- [12] T. Tuten, L. and M. Bosnjak, Understanding differences in web usage: The role of need for cognition and the five factor model of personality, *Social Behavior & Personality* 29 (2001) 391-398.
- [13] J. Paolillo, The virtual speech community: Social network and language variation on IRC, *Journal of Computer-Mediated Communication* 4 (1999) online: www.ascusc.org/jcmc/vol4/issue4/paolillo.html.
- [14] G. Riva and C. Galimberti, Interbrain frame: interaction and cognition in computer-mediated communication, *CyberPsychology & Behavior* 1 (1998) 295-310.
- [15] G. Mantovani, Social context in HCI: A new framework for mental models, cooperation and communication, *Cognitive Science* 20 (1996) 237-296.
- [16] F. Sudweeks and S. J. Simoff, Complementary Explorative Data Analysis: The reconciliation of quantitative and qualitative principles, in *Doing Internet research: Critical issues and methods for examining the Net*, S. Jones, Ed. Thousand Oaks, CA: Sage Publications, Inc., 1999, pp. 29-56.
- [17] D. M. Mertens, *Research methods in education and psychology*. Thousand Oaks, CA: SAGE Publications, 1998.
- [18] R. F. DeVellis, *Scale development: Theory and applications*. Newbury Park, CA: Sage Publishing, 1991.
- [19] P. A. Adler and P. Adler, Observational techniques, in *Handbook of qualitative research*, N. K. Denzin and Y. S. Lincoln, Eds. Thousand Oaks, CA: Sage Publishing, 1994, pp. 377-392.
- [20] J. P. Spradley, *Participant observation*. New York: Holt, Rinehart & Winston, 1980.

- [21] M. Q. Patton, *Qualitative evaluation and research methods*. Newbury Park, CA: Sage Publishing, 1990.
- [22] G. McCracken, *The long interview*. Newbury Park, CA: Sage Publishing, 1988.
- [23] D. L. Morgan, *Focus group as qualitative research*. Newbury Park, CA: Sage Publishing, 1988.
- [24] Y. S. Lincoln and E. G. Guba, *Naturalistic inquiry*. Newbury Park, CA: Sage Publishing, 1985.
- [25] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press, 1994.
- [26] L. C. Freeman, Visualizing Social Networks, *Journal of Social Structure* 1 (2000) online: <http://www.heinz.cmu.edu/project/INSNA/joss/vsn.html>.

EMERGING COMMUNICATION

Studies on New Technologies and Practices in Communication

EDITORS-IN-CHIEF: G. RIVA, PH.D., F. DAVIDE, PH.D.

[Http://www.emergingcommunication.com](http://www.emergingcommunication.com)

Volume 1: Communications Through Virtual Technologies: Identity, Community and Technology in the Communication Age

Editors: G. Riva and F. Davide

ISBN: 1 58603 162 7

Price: US\$90/EUR95/£60

Order: Copies:

Volume 2: CyberPsychology: Mind, Cognition and Society in the Internet Age

Editors: G. Riva and C. Galimberti

ISBN: 1 58603 197 X

Price: US\$95/EUR95/£60

Order: Copies:

Volume 3: Say not to Say: New perspectives on miscommunication

Editors: L. Anolli , R. Ciceri and G. Riva

ISBN: 1 58603 215 1

Price: US\$95/EUR95/£60

Order: Copies:

TOTAL AMOUNT:.....

EURO prices are definitive. In the USA and Canada the US\$ prices are definitive. Customers in EU countries please mention your V.A.T. number (MwST, TVA, BTW,) to receive an invoice without V.A.T. charge. Otherwise we will have to charge 6% V.A.T. For prepaid book orders we do not charge any postage or handling.

Name:.....
Organization:
Department:
Address: *ZIP/Postal Code:*.....
City: *Country:*
E-mail:..... *V.A.T. No:*.....

Please bill me (postage and handling charges will be added):

Charge my credit card: *American Express:* *Euro/Mastercard:* *Visa:*

Card No.:..... *Expiration date:*.....

Send the order form to one of the addresses below (by fax, traditional mail or e-mail)

IOS Press Nieuwe Hemweg 6B 1013 BG Amsterdam The Netherlands Tel: +31 20 688 3355 - Fax: +31 20 620 3419 market@iospress.nl	IOS Press, Inc. 5795-G Burke Centre Parkway Burke, VA 22015 U.S.A. Fax: +1 703 323 3668 iosbooks@iospress.com
--	--