CREATING HIGH-QUALITY PRODUCTS IN OPEN CONTENT VIRTUAL COMMUNITIES: A FUNCTIONAL GROUP PERSPECTIVE BASED ON THE TIME, INTERACTION, AND PERFORMANCE THEORY

Kévin D. Carillo

A Thesis
In
The John Molson School of Business

Presented in Partial Fulfillment of the Requirements for the Degree of Master of Science in Administration (DS/MIS) at
Concordia University
Montreal, Quebec, Canada

January 2006

Your file Votre référence
ISBN: 0-494-14363-0
Our file Notre référence
ISBN: 0-494-14363-0

ABSTRACT

Creating High-Quality Products in Open Content Virtual Communities: A Functional Group Perspective Based on the Time, Interaction, and Performance Theory

Kévin D. Carillo

This study examines the group input and group process factors that lead to high-quality products in open content communities. A theoretical model was developed, based on the Input-Process-Output approach of group research in social psychology and on the Time, Interaction, and Performance Theory proposed by McGrath (1991). The model tests the relationship between input variables: organizational (Organizational Support), group-related (Group Size, Shared Experience and Group Heterogeneity) and individual (Member Competency and Member Activeness); process variables: group production, group well-being and member support: and an output variable: group effectiveness. A quantitative study of the detailed activity logs of Wikipedia, an open content encyclopedia, was designed using a sample of 10,000 group tasks.

A partial least squared (PLS) statistical approach was used to test the theoretical model. This study shows the synergistic and overwhelming role played by group size and shared experience on both group process variables and group effectiveness. This project provides also evidence about (1) the positive effect of group heterogeneity on group production; (2) the positive effect of organizational support and member activeness on group well-being; (3) the positive effect of member activeness on member support; and finally, (4) the positive effect of organizational support and member activeness on group effectiveness.

ACKNOWLEDGEMENTS

I would like to thank my thesis supervisor, Professor Chitu Okoli, for his continuous support and guidance. He has been a true mentor who taught me the virtues necessary to quality academic research, which have even helped me beyond this master's thesis.

Special thanks to my committee members, Professor Anne Beaudry for her valuable feedback and for giving me some of her time out of her busy schedule, and Professor Bouchaib Bahli for his support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

> Collective intelligence [is the innovation that will most alter how we live in the next few years]. Think of how Wikipedia works, how Amazon harnesses user annotation on its site, the way photo-sharing sites like Flikr are bleeding out into other applications. ... We're entering an era in which software learns from its users and all of the users are connected.
>
> (Tim O'Reilly, Time Magazine, 2005)

As Tim O'Reilly, a pioneer and leader of the open content (OC) movement, claims above, the OC movement are enabling the advent of a major turn in information-based products through the enormous potential of collective intelligence. The OC movement relies on an alternate philosophical and sociological view of information product creation that posits the superiority of open systems relative to closed systems. It is an extension of the open source software (OSS) principles in the software development area, to all domains that pertain to information-based product development. The term "open content" has been defined as "content possible for others to improve and redistribute and/or content that is produced without any consideration of immediate financial reward—often collectively within a virtual community" (Cedergren, 2003). The application of such principles has already been successfully implemented not only in software development with renowned projects such as the Linux operating system, the Apache web server, or the Firefox internet browser, but also in various domains with projects such as Wikipedia (an open content encyclopedia), the Public Library of Science (online archives of peer-reviewed scientific articles), OpenCourseWare (open sharing of

undergraduate and graduate course content), Openlaw (experiment in the open crafting of legal arguments), and many others.

Nonetheless, the rise of the OC movement has occasioned questions about the reasons for creating high-quality projects. Even with several famous examples such as Linux that seriously competes with the giant Microsoft Windows, or Firefox which is making gains against Microsoft Internet Explorer, there is no common understanding of what makes open content practices so successful. This thesis introduces a new perspective to answering this question by adopting a group level approach based on the extensive past group research in social psychology and management information systems to investigate the factors that lead to high-quality open content performance.

1.1 About the open source and open content phenomena

The advent of the Internet continues to provide new promising ways to accomplish work tasks. Software development was among the first domains to be affected by this revolution. Interestingly, OSS (originally called "free software") appeared even before people started thinking in terms of proprietary software, as software development was originally guided by open source principles (von Hippel & von Krogh, 2003). In the 1960's and 1970's, software programming was mainly performed in both academic and corporate laboratories by scientists and engineers who freely gave, exchanged and modified software.

Open source communities believe in an alternative way to develop and distribute software, which has allowed them to effectively challenge and often outperform

proprietary software by enabling better reliability, lower costs, shorter development times, and a higher quality of code. The basic definition of open source software was expressed by the Open Source Initiative[1] in a document called the "Open Source Definition". This document articulates essential open source principles: free redistribution of both source code and compiled program; permission to create derivative works; integrity of the author's source code; no discrimination against persons or fields of endeavour; distribution of license that must not restrict other software; and must be technology-neutral (Perens, 1997).

Until today, several major OSS projects stand out in people's minds. The Apache web server project started in early 1995 and has become the most popular web server software, controlling over 66 percent of the market, much more than Microsoft and Netscape combined (Edwards, 1998; von Hippel & von Krogh, 2001). Inspired by Eric Raymond's paper "The Cathedral and the Bazaar" (Raymond, 2001), Netscape, one of the main actors in the Internet browser industry, decided to release the source code of its Communicator Internet client suite, thus creating the Mozilla project. In November 2004, the Mozilla Foundation released the Firefox 1.0 web browser. Nowadays, many major hardware and software actors have experimented with OSS such as IBM, Apple, HP, Microsoft, or else Google.

[1] Open Source Initiative (www.opensource.org) is an organization dedicated to promoting open source software. It was founded in February 1998 by Bruce Perens and Eric S. Raymond. .

1.2 From open source to open content philosophy

The OSS principles are not restricted to the software discipline; they are equally applicable to any other discipline that relies on creative intellectual work. As a result, the open source movement has expanded beyond the boundaries of software programming by giving birth to the broader notion of "open content". This term encompasses any type of creative information-based work including articles, pictures, audio, or video that is published under a license that explicitly allows the copying and editing of information. The original open content license is the GNU Free Documentation License, a copyleft license for free content, designed by the Free Software Foundation (FSF) ("GNU Free Documentation License", Wikipedia, 2005).

The first documented case of open content actually goes back to the seventeenth century, to the Royal Society in London which constantly promoted information sharing across the globe as a public enterprise ("Open content", Wikipedia, 2005). Nowadays, the OC revolution has given birth to many projects and initiatives that have been increasingly expanding due to their overall success. The Massachusetts Institute of Technology (MIT), a leader in both open source and open content, launched Open Courseware, a free and open educational resource for faculty and students throughout the world (www.ocw.mit.edu), initiating the OC movement in education. This field is the most represented area in terms of open content initiatives, with projects such as the Open Learning Initiative of Carnegie Mellon University, the Open Learning Support of Utah State University, the California Open Source Textbook Project University, and the Harvard University Library Open Collections Program. In 1998, OpenContent

(http://www.opencontent.org,) began evangelizing the idea of opening access to educational resources with the release of the Open Publication License and a small collection of educational resources. This website catalogs the major open content initiatives in education.

The Open Content Project was dedicated to creating open content. Primarily designed for academics, the project's Open Publication License was easily adaptable to the needs of the artist or other content provider. The Open Content Project has been succeeded by Creative Commons (www.creativecommons.org), a non-profit organization devoted to expanding the range of creative work available for others to legally build on and share, allowing copyright holders to grant some of their rights to the public while keeping others for themselves.

The OC trend has only recently begun to excite researchers' curiosity. Baldi et al. (2003) explored the MIT's courseware initiative by focusing on the promises of such a technical move in education. Very few articles have explored OC, the most complete one being written by Cedergren (2003), who proposed an OC-based value model based on an analysis of the driving forces for the cooperation between players that work with open content. The study focused on the dynamics of business development, technical design and legal aspects in this field.

As a result, the open source revolution has instigated the emergence of a new philosophy stream that relies on the free nature of any type of information and that emphasizes free collaboration and sharing among human beings as part of particular types of virtual communities.

1.3 Open content and open source communities

The scarcity of open content research means that the nature of OC communities not involving software development is not well understood. As a consequence, the only way to have some indications about OC members is to go back to the cradle of the "open" phenomenon. Open source software communities, a particular subset of OC communities, is a domain in which there has been a substantial amount of research. Open source software development communities have long intrigued researchers because of the apparent absence of structure and organization in OSS practices, as seen in Eric Raymond's depiction of the Linux community:

> No quiet, reverent cathedral-building here—rather, the Linux community seemed to resemble a great babbling bazaar of differing agendas and approaches ... out of which a coherent and stable system seemingly emerges only by a succession of miracles. (Raymond, 2001)

OS virtual communities are defined as groups of loosely connected programmers, who use the Internet as a medium for collaboratively developing, improving, and disseminating software (O'Reilly, 1999). OSS communities are a particular kind of virtual community (Carillo & Okoli, 2005; Chengalur-Smith & Sidorova, 2003; Diker & Scholl, 2001; Lakhani & von Hippel, 2003; Ljungberg, 2000). By referring to virtual communities, we adopt the following definition:

> A cyberspace supported by computer-based information technology, centered upon communication and interaction of participants to generate member-driven contents, resulting in a relationship being built up. (Lee et al., 2003)

The most commonly accepted view of the OS movement has been depicted by Eric Raymond in his famous paper "The Cathedral and the Bazaar"[2] (2001), where he characterizes it as a *gift* culture as opposed to the more common *exchange* culture. Raymond argues that the two cultures differ because an exchange culture relies on *scarcity* of resources whereas a gift culture is based on *abundance*. Through the redefinition of intellectual ownership, social status among OS communities is not determined by what you own but rather by what you give away (Hann *et al.*, 2002). Raymond (2001) posits that reputation (built from the amount of free contributions) and peer recognition mainly motivate people to freely and willingly participate in OSS projects.

Another interpretation has been suggested based on an economic perspective. Such a view posits that a person will contribute to an OSS development project only if the benefits of the contributions will outweigh its costs (Hann et al., 2002), which signifies that a contribution needs to generate a "net benefit" that consists of the sum of immediate payoff and delayed payoff (Lerner & Tirole, 2002). Lerner and Tirole identified two main immediate benefits: the increase in value of personal use of a product and the satisfaction of having achieved something valuable. The career future incentive (future jobs, shares, and access to capital market) and the ego gratification incentive (stemming from a desire of peer recognition) have been seen as the two main delayed benefits (Lerner & Tirole, 2002).

[2] The Cathedral and the Bazaar is an essay by Eric S. Raymond on software engineering methods, based on his observations of the Linux kernel development process and his experiences managing an open source project, fetchmail. It was first presented by the author at the Linux Kongress on May 27, 1997 and was published as part of a book of the same name in 1999. It is commonly regarded as the manifesto of the open source movement.

Research has definitely dispelled the stereotype of OSS developers as anarchistic hackers operating on the fringes of society (Carillo & Okoli, 2005; Fitzgerald, 2004). Several surveys have identified consistent traits among OSS developers such as a vast male majority, with an average age around 30. The majority of people are between 20 and 30, and 20% are students and over 50% have IT-related jobs. Most OSS developers have a bachelor's degree (or equivalent), a little less have a master's degree, and fewer than 10% received a PhD degree (Ghosh *et al.*, 2002; Hars & Ou, 2001; Lakhani *et al.*, 2002).

1.3.1 Research objectives

The open content perspective has become more than merely a viable method to produce creative and collective intellectual works. Until now, researchers have striven to understand the OC movement from a descriptive point of view. That is, most studies so far have focused on understanding OC communities in terms of ideologies, values, culture and participants' motives (Bergquist & Ljungberg, 2001; Ghosh et al., 2002; Hann et al., 2002; Hars & Ou, 2001; Lerner & Tirole, 2002). While acknowledging the overall success and quality of OC projects, the "bazaar-view" of the open source movement leads both researchers and practitioners to see OC as an obscure environment from which arise high quality projects. First, this view is rather simplistic in the sense that people tend to ignore that only a small proportion of OS projects have led to high-quality outputs. Moreover, even though opposed to the cathedral-like development method, bazaar-like interactions may still be governed by rules and standard group mechanisms among the

different members of such groups. Specific group patterns may characterize successful projects whereas others may be responsible for unsuccessful ones.

This research project focuses on answering the following research questions:

1. **What are the group input factors that contribute to high-quality products in open content practices?**

2. **What are the group process factors that lead to high-quality products open in content practices?**

Group research in MIS and social psychology has provided strong theoretically-based findings that may shed some light on open content groups functioning. Our study will help in identifying both group input and group process factors that affect OC project success. Furthermore, whereas group research has been criticized in its recurrent choice of artificial lab experiments (McGrath, 1991), this project will bring an important contribution to group research by using real data from live, functioning groups in an OC virtual community.

The content of this thesis is organized as follows: Chapter 2 will present a literature review of the Input-Process-Output perspective in group research in both social psychology and MIS. It will present and describe the different group input, process and output factors that have been studied in the literature, and then the Time, Interaction, and Performance theory (McGrath, 1991) will be introduced. In chapter 3, the proposed research model will be presented, and a set of research hypotheses will be developed. Chapter 4 will describe the chosen research methodology by introducing the selected

research design, the sample open content community, the operationalization of the research model constructs, the measurement instrument, and the data collection procedures. In chapter 5, the data will be analyzed using a partial least squared analysis (PLS) and a set of results will be presented. The last chapter will conclude with a detailed discussion of the findings, an explanation of the contributions and limitations, and future research propositions.

2 *LITERATURE REVIEW*

This study explores the open content project quality issue using a group perspective. Groups are defined as "complex, intact social systems that engage in multiple, interdependent functions, on multiple concurrent projects, while partially nested within, and loosely coupled to, surrounding systems" (McGrath, 1991).

This project focuses on group input and group process factors in relation high-quality group project in OC communities using two group-focused theoretical foundations: the Input-Process-Output approach of social psychology and the Time, Interaction, and Performance theory (McGrath, 1991). Both research streams will be presented in their original context of social psychology and management information systems (MIS).

2.1 Input-Process-Output models on group research

Input-Process-Output (IPO) models are the direct expression of the functional view of groups first introduced in group research in the domain of social psychology (Hackman, 1987; Hackman & Morris, 1975; McGrath, 1984; Steiner, 1972). According to Input-Process-Output models, both inputs to the group and the processes that groups use when working together influence whether groups will be effective; that is, whether they achieve their production goals, meet members' needs, and maintain themselves over time. Before introducing the IPO approach per se, it is important to understand its

foundations and basic assumptions inherited from the functional perspective. Those considerations are necessary, as they define the overall mindset adopted in this research project.

2.1.1 The functional perspective

The *functional perspective* examines groups in terms of the inputs and processes that function to influence group effectiveness (Poole *et al.*, 2004; Wittenbaum *et al.*, 2004). This perspective considers group performance as its main focus, which aligns well with our approach to tackling the question of performance of open content processes. The functional perspective is characterized by the three following assumptions (Hackman, 1987; Hackman & Morris, 1975; McGrath, 1984; Steiner, 1972):

1. **Groups are goal oriented**. The functional perspective assumes that groups have one or more goals that identify some purpose or goal to accomplish. These goals may be social-emotional (e.g., to provide support to members), group-oriented (e.g., to obtain resources for the group to continue), or task-oriented (e.g., to produce a product, idea, or decision). However, much of the research from the functional perspective has focused on the effective accomplishment of task-oriented goals. Task-oriented goals include production in the cognitive realm (e.g., reaching a good group decision, generating creative ideas, solving a problem, remembering information) or physical world (e.g., building widgets, performing a song).

2. **Group performance varies and can be evaluated.** According to the functional perspective, group performance is evaluated by some standard that indicates how well

the group reaches its goals. Most commonly, the standard consists of normative criteria that identify how groups should perform. These criteria are typically based on a rational model (e.g., members should communicate and optimally weigh all task-relevant information). When group performance falls short of this normative standard, interventions are generated to help groups reach their potential.

3. **Internal and external factors influence group performance via the interaction process.** The functional perspective presumes that factors emanating from within the group (e.g., member composition, group size) and external circumstances (e.g., outside threat, time pressure) affect how the group performs. Thus, group performance is a causal outcome of these internal and external inputs. Nonetheless, it does not only pertain to direct causation effects. A given input can lead to many different outcomes; likewise, one outcome can be produced by many different inputs. Moreover, inputs may interact with each other to produce a group outcome. The input-output relation is mediated by processes that occur during group interaction (e.g., communication patterns, conflict management). These interaction processes cause variations in group outcomes.

Functional researchers attempt to identify the group behaviours and activities that promote effective performance and also those that detract from it. Inputs that influence group functions include the nature of the group's task, the internal structure of the group, group cohesiveness, group composition, and the group's environment. Outputs considered in functional theories include group effectiveness as measured by productivity, efficiency, and quality, leadership effectiveness; and satisfaction with the group outcomes.

2.1.2 The general Input-Process-Output model

In Input-Process-Output (IPO) models, the inputs have both a direct effect on group outcomes and indirectly, by influencing the group process. Inputs include such resources as personnel, task, tools and time (Kraut, 2003). The basic IPO model was proposed by McGrath (1984) as seen in Figure 1. It posits that group input variables (categorized as individual, group and environment factors) influence group outputs (that are either group performance-oriented or not) through the mediation of group interaction process variables. McGrath is a leading developer of group research who criticized prior group research work as "context-stripped, ad hoc, laboratory groups of limited mission and conditions" (McGrath, 1991).

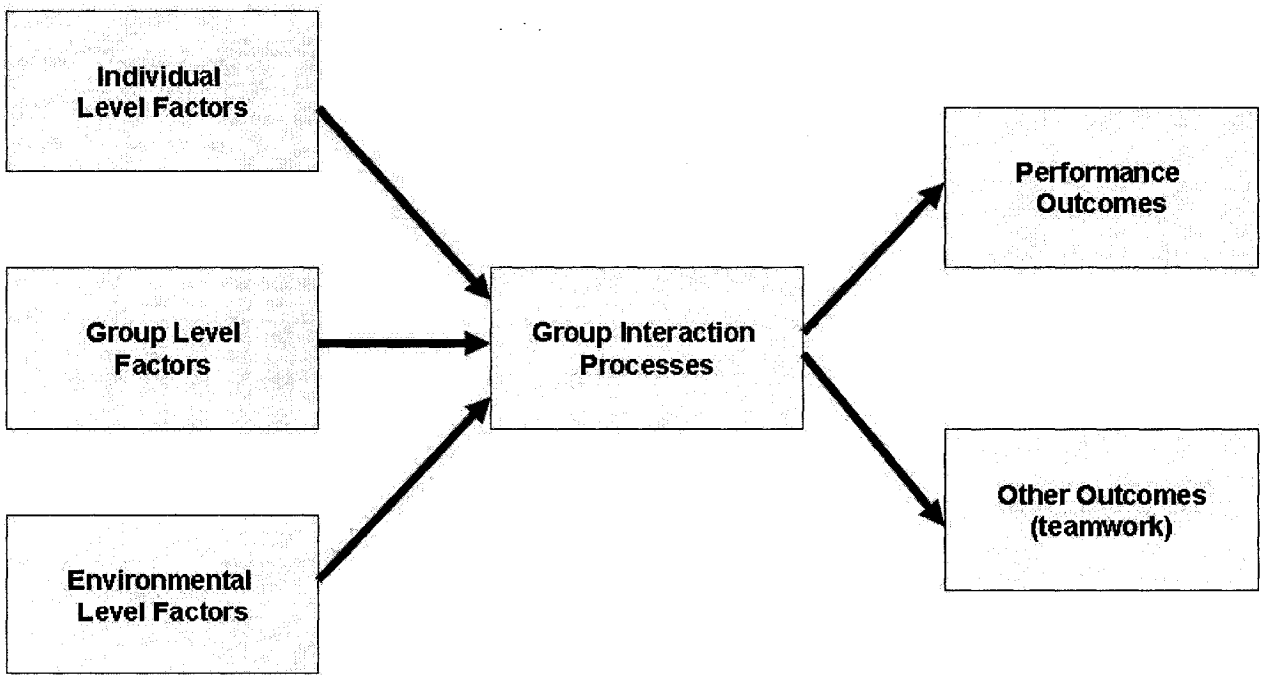

Figure 1: The Input-Process-Output Model (McGrath, 1984)

2.1.3 Input-Process-Output models in MIS research

Studies on group support systems (GSS) and group decision support systems (GDSS) have provided a vast array of well-grounded research. Inherited from social psychology, the functional perspective, and IPO models, GSS and GDSS research has proven very insightful in MIS.

2.1.3.1 Review of past MIS studies using an Input-Process-Output approach

Several early studies introduced the functional view of group research. DeSanctis and Gallupe (1987) have been very influential by providing one of the first frameworks for GSS research. They proposed a "contingency" theory to help explain the reason why GSS is not always as beneficial as people would think, depending on the nature of the technology and structure provided. Technology and structure are more or less appropriate depending on group size, task type and communication mode. Jelassi and Beauclair (1987) proposed a similar framework that addresses behavioural and technical aspects in order to enhance GSS impact.

Mostly based on the latter studies, research on IPO models has provided interesting results (Gallupe *et al.*, 1988; Pinsonneault & Kraemer, 1990). Pinsonneault and Kraemer clearly distinguish between group decision support systems (GDSS) and group communication support systems (GCSS), studying group processes and outcomes in both types of system. One the most active MIS researchers in group research, Alan Dennis, used an IPO-based model (Dennis *et al.*, 1988) to develop a model of Electronic Meeting Systems (EMS).

Nunnaker et al. (1991) further investigated EMSs, using both developmental and empirical research to introduce the Arizona EMS research program, whose aim was to produce group software and tools. The empirical part of their study mainly employed the IPO model and focused on the predictors of both process losses and process gains. As GSS research started providing deeper insights and findings, the functional approach was increasingly integrated in MIS research. Zigurs and Kozar (1994) developed a theoretical model of roles in computer-supported meetings that integrated the complete IPO approach. They examined the impact of a group support system on roles through a field study of 10 work teams. The study found a gap between the role expectations of meeting initiators and meeting participants and also a gap between participants' role expectations and actual roles filled. By studying how group attitudes and outcomes evolve over time with repeated use of a GSS, Chidambaram (1996) found that attitudes of GSS users changed over time from highly negative to somewhat positive, whereas outcomes improved more slowly.

As globalization has become a topical research issue, Tung and Turban (1998) proposed a research framework that addresses the major issues involved in the implementation of distributed GSS (DGSS) through a literature review. They used the IPO framework to classify the various concepts that were identified, based on Pinsonneault and Kraemer's (1990) framework. This paper is among the first extensive reviews that integrates all the major GSS issues highlighted by MIS research.

Other GSS research has been inspired by the functional view offered by social psychology. Burke and Chidambaram (1999) conducted a longitudinal study of media characteristics and group outcomes. They found that face-to-face groups initially

considered that their medium was warmer, had a better interface, and was more effective compared to the distributed option. However, after a certain period of time no significant difference was found. It was also found that synchronous and asynchronous means of communication do not differ in their perceptions or performance.

In order to test whether the use of a GSS had a positive effect on the quality of group processes and outcomes, Batenburg and Bongers (2001) combined several forms of evaluation in a field experiment. They found that process facilitation and time effects had more influence on group processes and outcomes than did the use of GSS. Lurey and Raisinghani (2001) explored effectiveness within virtual teams. In an effort to determine the factors that contribute to or inhibit the success of a virtual team, a survey was distributed to eight companies in the high technology, agriculture, and professional services industries. The results were integrated using the IPO model. It was found that teams' processes and team members' relationships with each other presented the strongest relationships to team performance and team member satisfaction, while the selection procedures and executive leadership styles also exhibited moderate associations to these measures of effectiveness.

In addition, by studying the influence of communication mode and incentive structure on GDSS process and outcomes, Barkhi et al. (2004) compared the decision process and outcomes of groups that use a face-to-face GDSS to those that use a distributed GDSS operating under two different incentive structures. They found that communication mode and incentive structure can influence the effects of each other.

2.1.3.2 Fjermestad and Hiltz' review (1999)

The most extensive review thus far on processes and outcomes in computer-supported group decision making was conducted by Fjermestad and Hiltz (1999). This review integrates 200 different controlled experiments that were published up to mid-1998 in 230 articles in referred journals or major conference proceedings.

The paper categorizes all reviewed variables: systems, independent, intervening, adaptation and dependent variables, and analyzes 1,582 hypotheses resulting from pairings of independent and dependent variables. The framework that was used to integrate such an enormous bulk of MIS research is inspired from several studies that are based on the IPO approach (Dennis *et al.*, 1988; DeSanctis & Gallupe, 1987; Hiltz *et al.*, 1991; Jelassi & Beauclair, 1987; Nunamaker *et al.*, 1991; Pinsonneault & Kraemer, 1990; Poole & DeSanctis, 1990). It consists of four major categories: contextual or independent variables, intervening variables, group adaptation process variables, and outcome variables. Contextual variables are classified in line with prior IPO model-based studies by subcategorizing them into technological, group, task and context variables. Intervening factors are more related to methodological issues, whereas adaptation factors characterize group processes. Finally, outcome factors refer to effectiveness, efficiency, satisfaction, consensus and usability variables.

The assessment of OC group performance may be performed through different possible dependent variables based on Fjermestad and Hiltz' classification. Among the categories stated above, two are particularly related to the type of group system that is used: consensus (decision agreement, commitment) that particularly concerns GDSS

(Huang et al., 2003; Limayem & DeSanctis, 2000; Shirani et al., 1998); and usability measures that pertain to technological considerations (Alonzo & Aiken, 2004; Batenburg & Bongers, 2001). These do not appear relevant when applied to the performance of OC processes. Another category concerns satisfaction measures such as participation (Paul et al., 2004), cohesiveness (Huang et al., 2003), conflict management (as an outcome), influence (Huang et al., 1999), or confidence; these are all perception-related notions. Such measures do not capture any aspects of group tasks. Finally, the two most studied issues are efficiency and effectiveness measures. Depending on technology and task type, the efficiency measures that have been used are varied. They include decision time (Limayem & DeSanctis, 2000; Mathiyalakan, 2002), number of decision cycles, number of ideas (Dennis *et al.*, 1999; Huang & Wei, 1997; Parent *et al.*, 2000; Wong & Aiken, 2003), time spent in activities, and time spent waiting for responses for instance.

Effectiveness measures have varied widely in group research. Several approaches have been used: communication (Kayworth & Leidner, 2001), number of comments (Easton *et al.*, 2003), idea quality (Dennis *et al.*, 1999; Parent *et al.*, 2000), decision quality (Barkhi *et al.*, 1999; Huang & Wei, 1997; Paul et al., 2004; Saunders & Miranda, 1998), decision confidence (Limayem & DeSanctis, 2000), process quality, creativity or innovation, level of understanding (Kwok & Khalifa, 1998), task focus, depth of evaluation (Limayem & DeSanctis, 2000), and commitment to results (Huang et al., 2003).

2.2 McGrath's Time, Interaction, Performance (TIP) theory (1991)

Through the analysis of information sharing among group members, McGrath inferred after several experiments that in all the studied groups, unplanned change had occurred, which he mainly attributed to temporal pacing. Instead of focusing on a predetermined sequence of events in a group's life as groups fulfill their objectives and purposes (which was the prior academic approach), McGrath proposed a new view of group processes by focusing on the underlying contextual factors that are responsible for the observed shifts in group activities. McGrath (1991) and later Arrow and McGrath (1995) proposed theories in order to explain unplanned change. Building on the concept of social entrainment, introduced by Kelly and McGrath (1985) then by Kelly (1988), as the proposed cause of such change, McGrath proposed a new perspective of group process in his Time, Interaction, and Performance theory.

2.2.1 Main assumptions of TIP theory

In 1991, McGrath introduced the Time, Interaction and Performance (TIP) theory that posits that group coordination behaviours occur at different levels: within individual members themselves, among team members, and within the social context in which they operate. McGrath's TIP theory states that groups simultaneously perform a number of tasks that belong to three different levels: production, well-being and member support in relation to a specific group objective. In other words, groups contribute to "the systems in

which they are embedded, their component parts, that is, their members and to the group itself, as an intact and continuing social structure" (McGrath, 1991). The three group activity levels are:

1. **Group Production Function**: The relation between the group as a functional entity and the environmental conditions and constraints within which a group operates. It thus refers to the set of activities that a group performs to work on a common task.

2. **Group Well-being Function:** The activities that are related to the development and maintenance of a group as a system; hence, the relations among group members.

3. **Member Support Function:** The activities that are in relation to the ways an individual is embedded within a group that thus reflects the relations between individual members and the group.

2.2.2 TIP theory in MIS research

As MIS group research has focused on designing artificial groups through laboratory experiments, a temporal view of group processes such as is presented by the TIP theory was not applicable. Such experiments were mostly executed in short periods of time with tasks designed by researchers beforehand. As a result, the notion of group formation and membership was completely ignored. However, an exception is found in Dennis and Reinicke (2004). Using an analogy with the VHS vs. Sony Beta videocassette competition, they argued that brainstorming sessions may not be primarily concerned with the number of ideas generated when planning a brainstorming session, but rather may equally seek group well-being and member support. The authors developed

theoretical arguments and empirical evidence that suggest that electronic brainstorming is not as effective as verbal brainstorming at providing group well-being and member support. Other group-focused MIS studies have referred to TIP theory to provide insights to their findings, but very rarely was such theory used, operationalized, and tested (Dennis & Garfield, 2003; Malhotra et al., 2001; Massey et al., 2002; Shim et al., 2002). Moreover, a review of both social psychology and MIS research that emphasized the role of processes in group research used TIP theory to develop a conceptual model that focuses on task and social interaction using an influence perspective (Huang & Wei, 2000).

In conclusion, the social notion of group has emerged from the need for understanding OC practice overall success. Such view enables researchers to more comprehensively understand OC processes by focusing on a level of analysis (that is, group tasks) that will unveil the factors that contribute to OC product quality. Relying on the previous theoretical foundations taken from group research in social psychology and MIS and particularly the TIP theory, a theoretical model will be developed in the following chapter to help answer the research questions.

3 *RESEARCH MODEL AND HYPOTHESES*

This chapter details the development of the theoretical model that focuses on the group input and group process factors that lead to high-quality products in OC communities. It integrates the Time, Interaction and Performance theory (McGrath, 1991) to conceptualize group processes. First, the theoretical model will be presented, and then the research hypotheses will be developed.

3.1 Research model

In order to have an overall picture of the functioning of OC groups, this study encompasses both the functional perspective of group research in its IPO form and the Time, Interaction and Performance Theory (McGrath, 1991) by focusing on group factors that best explain group effectiveness. Such an approach is in line with McGrath's view of group research where he emphasizes the need for considering both contextual and temporal criteria in the study of group effectiveness. Thus, three types of variables will be considered: input variables, group process variables and output variables. The group input variables consist of environmental factors (Organizational Support), group factors (Group Size, Shared Experience, and Group Heterogeneity), and individual factors (Member Competency and Member Activeness). Furthermore, group process variables will be in line with TIP by being categorized among the three group functions of TIP: Group Production, Group Well-being, and Member Support. Each input variable is hypothesized to be positively related to each of the process variables. Similarly, the three

process variables are hypothesized to have a positive relationship with group effectiveness. The overall model that will be used in this study is displayed in Figure 2.

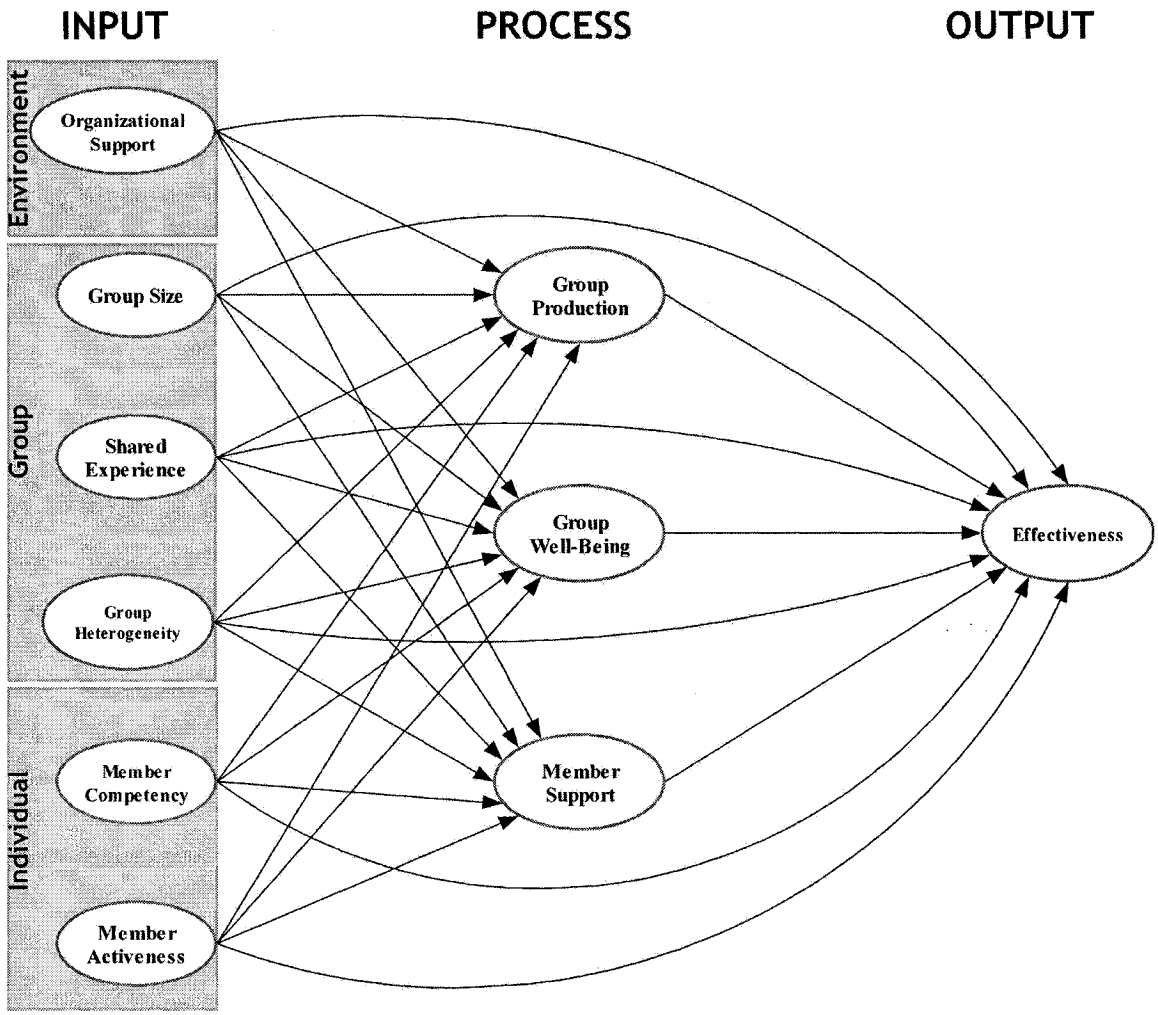

Figure 2: Model of group processes in open content communities

3.2 Research hypotheses

Group effectiveness is defined as the extent to which a group is able to perform a certain group task that fulfills a pre-determined list of quality and excellence standards.

Group effectiveness was chosen as the primary dependent variable for it is particularly appropriate in the case of information-based task production where quality can be easily measured through the categorization and specification of quality standards in OC communities.

3.2.1 Input variables

Three main categories of input factors will be studied: factors that concern the entire group; individual member factors; and organizational factors that take into consideration the contextual nature of group work.

3.2.1.1 Context variables

The management literature has provided results that indicate that in an organizational context, perceived organizational support leads to an increase in employees' creativity (Scott & Bruce, 1994). Organizational Support is defined as "individuals' perception that a relevant organization values their contribution in the context of group tasks" (LaMastro, 2001). MIS research has also emphasized the importance of organizational support in software simulation (McHaney & Cronan, 2000) and microcomputer usage (Anakwe *et al.*, 1999). In the OC context, where innovativeness and creativity have been long prevalent, organizational support seems to play an important role by emphasizing the influence of an OC community on its functioning and processes. This leads to the following hypotheses:

Hypothesis 1a: Organizational support is positively related to the group process variables: (i) group production; (ii) group well-being; and (iii) member support.

Hypothesis 1b: Organizational support is positively related to group effectiveness.

3.2.1.2 Group composition variables

Group composition variables characterize groups as discrete entities. Three variables will be included in the designed model: group size, shared experience and group heterogeneity.

Group Size

Group Size is the number of people who have contributed to achieving a specific group task. Group size has been one of the most explored issues in social psychology, management and MIS when studying groups (see Fjermestad & Hiltz, 1999, for a complete summary of the results). For instance, prior studies have shown that a minimum group size of about eight is necessary for successful use of a GSS (Aiken *et al.*, 1994). Moreover, in a meta-analysis about the impact of the use of GSS on groups, it was also found that group size may be an important moderator when measuring decision time and satisfaction with the process, with decision time being shorter and satisfaction higher for larger groups (Dennis & Wixom, 2001). Furthermore, a study of electronic meeting systems (EMS) using two concurrent experiments with groups of varying size analyzed the number and quality of unique ideas generated by groups of each size using electronic

and non-electronic verbal brainstorming. It was found that larger groups generated more unique ideas and more high-quality ideas, while members were more satisfied with the EMS (Gallupe *et al.*, 1992). Therefore, the following hypotheses can be suggested:

Hypothesis 2a: Group size is positively related to the group process variables: (i) group production; (ii) group well-being; and (iii) member support.

Hypothesis 2b: Group size is positively related to group effectiveness.

Shared Experience

Shared Experience is the extent to which the group members of a particular group task have previously collaborated in other group projects. In Fjermestad and Hiltz's (1999) study, 25 articles were found to have group composition variables as independent variables. They found that only eleven articles (out of 200) used groups that were established before the study whereas the rest used groups formed only for the purpose of the experiment, of which the majority used students as group members. As a result, there is a lack of investigation of the shared experience issue in relation to group work performance; we posit that shared experience may be an important factor in explaining both group processes and outcomes. Thus, the following hypotheses are drawn:

Hypothesis 3a: Shared experience is positively related to the group process variables: (i) group production; (ii) group well-being; and (iii) member support.

Hypothesis 3b: Shared experience is positively related to group effectiveness.

Group Heterogeneity

Group Heterogeneity is defined as the extent to which group members have varying characteristics such as the diversity of group activity, personality traits, attitudes, backgrounds, and abilities (Goodman *et al.*, 1986). Heterogeneous groups may enjoy a wider base of experiences, skills, abilities and perspectives that can help groups to be more effective in group tasks (Shaw, 1981). Such diversity in membership expands the resources available to the group, thereby increasing the likelihood of improved productivity (Shaw, 1981). It has been demonstrated that group heterogeneity can lead to performance gains by improving the quality of the decision strategies employed by workgroups (Steiner, 1972). In a study that explored the relationship between group heterogeneity, group rewards and success participation in system development, insightful results were found (Aladwani *et al.*, 2000). The authors concluded that both group heterogeneity and group-based rewards impact participation, and as participation positively influences outcomes, both variables indirectly impact outcomes.

Holsapple & Luo (1999) examined the effects of different group-work patterns in a GSS environment varying with group members' experience or ability levels. The authors found that experience-based work patterns (experienced participants working on a problem first and then passing their results on to less experienced participants) appear to have a positive impact on group decision quality when compared to the conventional work pattern under GSS environment (people of differing experience levels working simultaneously) .Therefore, the following hypotheses were derived:

Hypothesis 4a: Group heterogeneity is positively related to the group process variables: (i) group production; (ii) group well-being; and (iii) member support.

Hypothesis 4b: Group heterogeneity is positively related to group effectiveness.

3.2.1.3 Member characteristics

Member characteristics potentially include any attributes of individual members such as their attitudes, personality traits, age or previous experience with systems and tasks, as well as group members' activeness in the community. In order to study the influence of member characteristics in a group-focused research project, each group has to be described in terms of the average of all chosen individual characteristics among the group members.

Fjermestad and Hiltz found that both job tenure and member experience have been important issues that have been studied so far in GSS research. Member experience is the extent to which an individual has participated in group projects in an organization. Member tenure is the amount of time spent by an individual in an organization or community. Both issues deserve some investigation.

Furthermore, in a study of computer-supported meeting system and the influence of facilitator characteristics, it was found that the amount of experience and training, the amount of external versus internal facilitation, and use of GSS all correlate with multiple aspects of pre-meeting planning and agenda use (Niederman & Volkema, 1999). Member experience can also be associated with members' tenure in an organization. For instance,

Kelly & Bostrom (1998) related facilitator tenure with how facilitators manage socio-emotional issues in a GSS environment. The results showed that experienced facilitators are more inclined to deal with socio-emotional issues. We posit that, in a virtual community, member experience may also be categorized in terms of position nomination such as administrative positions or peer recognition seen as a direct consequence of experience in a community.

It is thus apparent that the member experience issue encompasses broad and varied notions, thus deserving a further sub-categorization. In this study, we distinguish between two general categories of member experiences. Member Competency refers to all member characteristics that describe a member's innate skills, capabilities, knowledge and roles that may be valuable in fulfilling group tasks. This is distinct from Member Activeness, which refers to the extent to which a member has actively participated in group tasks in an organization. The following hypotheses can thus be stated:

Hypothesis 5a: Member competency is positively related to the group process variables: (i) group production; (ii) group well-being; and (iii) member support.

Hypothesis 5b: Member competency is positively related to group effectiveness.

Hypothesis 6a: Member activeness is positively related to the group process variables: (i) group production; (ii) group well-being; and (iii) member support.

Hypothesis 6b: Member activeness is positively related to group effectiveness.

3.2.2 Group process variables

The Time, Interaction and Performance Theory framework (McGrath, 1991) was used to categorize process variables among the three interaction levels: group production, group well-being, and member support.

3.2.2.1 Group Production function

Group Production refers to the set of activities that a group performs, seen as a functional entity, to work on a common task (McGrath, 1991). The group production function is the system in which task performance occurs effectively. Groups are thus seen as functional units whose output is the task that is effectively done.

Several aspects of the group production function have been investigated in past studies. Participation was explored by Pinsonneault and Kraemer (1990) whose review of the empirical research suggests that GSSs increase participation of participants of electronic groups. Furthermore, Fjermestad and Hiltz (1999) identified six studies that investigated participation equality as a process variable, whereas seven focused on level of effort. In another review whose purpose was to unveil the differences between electronically-supported GDSS to electronically-supported and face-to-face meetings, equality of participation did not differ (Tung & Turban, 1998). A longitudinal study that compared the developmental patterns of groups in three different electronically-supported modes (face-to-face, dispersed asynchronous and dispersed-synchronous) found no particular difference among the three modes (Burke & Chidambaram, 1994). Furthermore, another project found that greater inhibition occurs in face-to-face groups

leading to the potential for less equality of participation than in distributed EMS groups while more equal participation was observed (Kutsko & Smith, 1991). As a consequence, the following hypothesis is proposed:

Hypothesis 7: Group production is positively related to group effectiveness.

3.2.2.2 Group Well-being function

Group Well-being describes the activities that have to do with development and maintenance of a group as a system (McGrath, 1991). The group well-being function consists of all actions that make contributions to the group itself as a distinct and continuing social structure, such as roles assumed by some members and behavioural norms (Dennis & Reinicke, 2004; McGrath, 1991).

Past research has studied several aspects of the group well-being function. Communication behaviour was one of the issues that has been explored (Adkins *et al.*, 2003; Pollard, 2003). Fjermestad and Hiltz (1999) also found twelve studies that addressed communication as a process variable. The level of information exchanged during group processes has also appeared as an important factor. Fjermestad and Hiltz (1999) found seven articles about information exchange, three about information credibility and ten about information sharing.

In a study intended to analyze group processes through an extended review, Huang and Wei (2000) emphasized the role of task interaction. Other group well-being issues such as coordination have been identified by the literature (Fjermestad & Hiltz, 1999). In a research project that examined the effect of system restrictiveness of

coordination structures in an asynchronous environment, it was concluded that groups with parallel coordination mode have a stronger belief that the decisions they made are of higher quality than those of groups with sequential coordination mode (Kim *et al.*, 2002). The management literature also provided interesting results about coordination. A study that explored the effects of temporal coordination on virtual teams in the context of Lotus Notes, showed through an experiment that coordination has moderating effects on group outcomes (Montoya-Weiss *et al.*, 2001).

Group conflicts have raised many problems in the practical world, because of cultural differences among group members and because of team heterogeneity in general (Oetzel, 2001; Paul et al., 2004). Conflicts are part of the group well-being function for it characterizes a certain type of group member interaction. Group conflict was studied by Bose and Paradice (1999) in an investigation of cognitive conflicts associated with the use of level 1 and level 2 GDSSs (DeSanctis & Gallupe, 1987). The results suggest that GDSSs reduced disagreement between group members and improved consistency of judgments better than other meeting environments would. From this body of research, the following hypothesis is proposed:

Hypothesis 8: Group well-being is positively related to group effectiveness.

3.2.2.3 Member Support function

Member support refers to the activities that have to do with the ways in which the individual is embedded within the group; it hence describes the relations between individual members and the group itself (McGrath, 1991). The lack of both temporal and

contextual considerations in group research has somehow caused member support issues to be neglected. The member support function consists of actions that make contributions to the individual members themselves (Dennis & Reinicke, 2004). It concerns all process issues that are related to member compensation, paybacks, and relationship building mechanisms. Compensation and payback considerations may concern either direct rewards (financial and other incentives) that may be granted by an organization or any knowledge or skills that an individual member may gain through group work.

In a study that investigated idea generation in the context of brainstorming, three types of groups were compared: verbal groups, nominal groups (each member working separately without being aware of the other group members contributions), and electronic groups. Based on TIP theory, the authors operationalized the member support function as both status auction (referring to the way employees establish their status in terms of their abilities and skills relative to other group members) and knowledge network (Dennis & Reinicke, 2004). They concluded that electronic brainstorming is less effective at providing member support.

A study that has linked incentive structures and performance inferred that group performance such as task performance, consensus, process satisfaction and participation were all higher with group-based incentives, whereas negative incentives led to more participation and less satisfaction (Shirani et al., 1998). Another study explored communication mode and incentive structure (Barkhi et al., 2004) to discover how these two factors influence group decision making. Results indicate that communication mode and incentive structure can influence the effects of each other. The authors thus

concluded that the appropriate design of incentive structures may be important to the success of virtual organizations (Barkhi et al., 2004).

In addition, McGrath (1991) highlights relationship building as one of the main aspects of member support. A study of the way in which leaders develop relationships with their virtual team members found that the leaders considered it essential to build some level of personal relationship with their virtual team members before commencing a virtual working relationship (Pauleen, 2003). As a consequence, the following hypothesis was tested:

Hypothesis 9: Member support is positively related to group effectiveness.

3.2.3 Summary of hypotheses

The development of the model has leaded to the following set of hypotheses as seen in Table 1.

Table 1: Research hypotheses

Hypothesis Label	*Content*
H_{1a}	Organizational Support is positively related to the group process variables: (i) Group Production; (ii) Group Well-being; and (iii) Member Support.
H_{1b}	Organizational support is positively related to Group Effectiveness.
H_{2a}	Group Size is positively related to the group process variables: (i) Group Production; (ii) Group Well-being; and (iii) Member Support.
H_{2b}	Group Size is positively related to Group Effectiveness.
H_{3a}	Shared Experience is positively related to the group process variables: (i) Group Production; (ii) Group Well-being; and (iii) Member Support.
H_{3b}	Shared Experience is positively related to Group Effectiveness.
H_{4a}	Group Heterogeneity is positively related to the group process variables: (i) Group Production; (ii) Group Well-being; and (iii) Member Support.
H_{4b}	Group Heterogeneity is positively related to Group Effectiveness.
H_{5a}	Member Experience is positively related to the group process variables: (i) Group Production; (ii) Group Well-being; and (iii) Member Support.
H_{5b}	Member Experience is positively related to Group Effectiveness.
H_{6a}	Member Activeness is positively related to the group process variables: (i) Group Production; (ii) Group Well-being; and (iii) Member Support.
H_{6b}	Member Activeness is positively related to Group Effectiveness.
H_{7}	Group Production is positively related to Group Effectiveness.
H_{8}	Group Well-being is positively related to Group Effectiveness.
H_{9}	Member Support is positively related to Group Effectiveness.

4 *METHODOLOGY*

This chapter introduces the methodology that will be used to test the proposed theoretical model. The first section presents the chosen research design and highlights why a quantitative field study of a virtual community is particularly relevant in contrast to the predominance of artificial experiments in past research. The second section presents Wikipedia, an open-content web-based encyclopedia that will be used to test the theoretical model. Finally, the operationalization of the group input, group process and group output variables is explained. The study of the Wikipedia open content (OC) community is expected to provide a clear understanding of OC group functioning in light of TIP theory.

4.1 Research design

McGrath has decried the lack of contextual and temporal considerations in group research whereas both dimensions play such an important role: "A very large proportion of past research on small groups has been done on a one-shot basis—study of groups newly formed for purposes of research, with the study extending only for a short interval during a single interaction" (McGrath et al., 1993, p. 415).

Moreover, group research has also been criticized in its recurring choice for using lab experiments owing to the artificial nature of the tasks that were assigned to the studied groups (Arrow & McGrath, 1995; McGrath, 1984, 1991). As a consequence, a

quantitative field study of a virtual community appears to be the most appropriate research design for this study, for two main reasons. First, it will allow the study of groups in their native context of a real OC community, which will respect context of their operation. Second, it will allow the consideration of tasks that are already known and natural to groups, which will minimize bias in the results. Indeed, the focus of this research is not related to the type of tasks achieved by groups as classified by McGrath (1984), but rather to study the overall behaviour of groups that are driven by an open content philosophy and culture. Thus, there is no need for controlling the tasks performed by groups.

A second issue that answers McGrath's view of group research in terms of lack of temporal considerations is made possible by OC technologies. OC platforms use the Internet as a means of communication and interaction, which allows tracking and storing exchanges that occur among group members. Such voluminous stored data allows the tracking of group functioning through time for all group process levels identified by the TIP theory classification. It is thus possible to have access to the temporal evolution of group processes with details for each individual information transaction.

4.2 Research sample

To test the model introduced in the previous chapter, it was decided to investigate Wikipedia (http://en.wikipedia.org), the largest open content encyclopedia community. Wikipedia is a web-based, multi-language, free-content encyclopedia written collaboratively by volunteers and sponsored by the non-profit Wikimedia Foundation.

As of January 2006, Wikipedia contains over 3.1 million articles in all its 200 languages (around 100 are active). More than 934,000 of these are in English, more than 344,000 in German, and more than 228,000 each in French. In each language edition, the number of articles has increased exponentially since Wikipedia's creation in 2001. Language editions operate independently of one another and are considered distinct encyclopedias. As the English version is by far the most mature and developed in terms of number of articles, it was decided to focus only on the English version for our study. A daily average of 60 million pages is requested.

Any visitor may edit Wikipedia's articles and have their changes instantly displayed. Wikipedia participants (called "wikipedians") rely on the principle that collaboration among users will significantly improve article quality over time—the same principle that guides open source software development, encapsulated in Linus' Law (Neus, 2001): "Given enough eyeballs, all bugs are shallow". Wikipedia's contributors do not require any particular expertise or formal skills in order to be entitled the right to edit an article about a specific subject. They are only required to adopt a "neutral point of view" and to include only well-documented facts—Wikipedia is an encyclopedia of accepted knowledge, not of original ideas. Moreover, according to the terms of the Free Documentation License under which Wikipedia content is licensed, users are clearly warned that their contributions may be "edited mercilessly and redistributed at will" by anyone ("Wikipedia", Wikipedia, 2005). Articles are not controlled by any particular user or editorial entity, and decision-making about conflicting or controversial issues in articles is carried out by democratic vote of Wikipedia contributors.

In January 2006, the English Wikipedia community had more than 43,500 active contributors (wikipedians who have contributed at least ten times since they arrived) among 840,000 registered users. On average, more than 3,000 new users register each month. Furthermore, more than 16,000 wikipedians make five or more contributions per month, while more than 2,000 make 100 or more contributions per month. Every day, around 1,700 articles are created and each article has an average of between 24 and 25 contributions. 74% of the articles have a size equal or larger than 85 words whereas around 32% have a size equal or larger than 341 words.

For this research study, on April 21, 2005, the entire Wikipedia English database in MySQL format was downloaded and installed—it is freely available, being open content. At that time, Wikipedia contained a total of 545,486 articles, with a total database size of around 40 gigabytes. Because of computing and memory limitations, a sampling policy was employed to select just 10,000 articles.

The output variable for the model was operationalized as "featured" articles, that is, those elected by community members as being of the highest quality that Wikipedia had produced (only 1 in 900 articles meet this standard). Featured articles will be explained in more detail in the operationalization section below. Only articles whose size was equal to or greater than 5,000 characters (833 words) were selected for the study sample. Since the featured articles were all larger than this size (all of the 580 featured articles were over 7,000 characters, except two that were 5,601 and 5,832 characters respectively), this threshold was chosen with the assumption that an article needs to be at least as long as the shortest featured article to possibly reach this quality standard. This left 58,169 articles. However, due to computational constraints, only 10,000 articles were

retained for this study. All the featured articles, and all articles above 5,000 characters nominated but not selected for featured status, were included in the sample, accounting for 1,073 articles. In addition, a random sample of 8,927 of other articles was selected, for a total of exactly 10,000 used for the study.

In addition to the 10,000 selected articles, the corresponding talk pages were also downloaded. Talk pages allow contributors to interact, discuss and debate issues in direct relation to the article itself. Furthermore, the user pages and user talk pages of all the users having contributed at least once in one of the 10,000 articles were downloaded. User pages deal with Wikipedia users' personal presentations and auxiliary pages for personal use whereas user talk pages are used to leave messages for users and to personally interact with them.

4.3 Operationalization and measurement

This section introduces the operationalization of each variable by defining the measurement items that will be used to assess each theoretical construct.

4.3.1 Output variable

The effectiveness of the article production process deals with overall article quality. To estimate this notion, we devised a scoring system based on whether the article had been elected as a featured article. As earlier described, a featured article is a particularly well-written and complete Wikipedia article, which meets several key quality criteria.

It has to be comprehensive, accurate, stable, well-written, uncontroversial in its neutrality and factual accuracy, and has to comply with the style standards required by Wikipedia, have relevant images and an appropriate length.

Each Wikipedia language version independently maintains its own featured or high-quality article selection process. In the English Wikipedia, each day, a high-quality article is chosen and displayed on the front page of Wikipedia, hence the name of "featured article". Anybody may nominate an article as a potential featured article. A rough consensus has to be reached for an article to be promoted to "featured" status. If sufficient time has elapsed without objections being resolved, the nomination will be rejected. On the date of the sample download (April 21st, 2005), there were 580 featured articles. In addition, as of that date, there were 497 articles that had been nominated for featured status but had been rejected. In this study, we refer to these articles as "nominees". Thus, three article quality levels can be identified, and will constitute our article quality measure:

- Regular article, with no nomination for featured status selection. This counts for a score of "0".
- Article nominated for featured status, but not accepted. This counts as "1", assuming that the article was of some quality higher than a regular article in order to have merited nomination.
- Article nominated and elected for featured status. This counts as "2".

4.3.2 Input variables

The operationalization of the input group factors that were previously introduced in the model development chapter is presented in the following sections.

4.3.2.1 Organizational support

The Wikipedia community constantly encourages its members to improve articles to featured article status. Various mechanisms exist for members to draw attention to particular articles so that other members can integrate the comments and select which articles deserve a concerted contribution effort from the community, the main aim being to improve selected articles to reach sufficient quality to be nominated for featured status. The English Wikipedia encourages its contributors to participate in contributing to certain articles using four different strategies:

1. Collaboration of the week[3]: Each week wikipedians vote on a topic for a new article that did not previously exist, whose topic appears as particularly relevant.

2. Article Improvement Drive[4]: This is similar to Collaboration of the week, but this category is for articles that already exist, but are judged not to have a sufficient degree of completeness and quality.

3. Cleanup[5]: Some articles may have problems in terms of grammar, structure, or be otherwise somehow confusing. A special request may be posted by anyone for any article so as to provide further work.

[3] http://en.wikipedia.org/wiki/Wikipedia:Collaboration_of_the_week
[4] http://en.wikipedia.org/wiki/Wikipedia:Article_improvement_drive

4. Articles needing attention[6]: Certain articles need special attention from people familiar with specific topics. Expert wikipedians are sometimes needed in order to complete the writing of an article.

Another avenue for drawing attention to an article is called "Peer Review"[7]: Contributors can request peer reviews for certain articles that deserve quality review and input from other wikipedians on how to steer the development of the article, with the explicit intention of eventual nomination for featured status. Due to the overall importance of peer review requests as opposed to the other organizational support strategies, the organizational support score of an article will be measured by two distinct items:

- Number of times a peer review was requested for an article.
- Number of times collaboration of the week, article improvement drive, cleanup, and articles needing attention requests were formulated in regard to an article.

4.3.2.2 Group size

As described in the literature review, group size has been one of the most studied factors in investigating group performance. Group size will be measured by the number of non-anonymous contributors that have directly contributed any content to an article. While anonymous contributors are an important part of the Wikipedia community, this

---

[5] http://en.wikipedia.org/wiki/Cleanup
[6] http://en.wikipedia.org/wiki/Articles_needing_attention
[7] http://en.wikipedia.org/wiki/Wikipedia:Peer_review

study only considers actions by traceable users who have participated while logged in to Wikipedia user accounts.

4.3.2.3 Shared experience

Shared experience is purely task-related in the context of this study, which means that it refers to the accumulated experience of group members during the article production process. Shared experience will be measured with two items:

- The number of times that pairs of contributors to the article have interacted in writing any other Wikipedia article, divided by group size. The choice for dividing this number by group size is due to the expected high correlation that may arise with group size (which is another construct of the model). It was thus decided to measure shared experience as "the average amount of shared experience per contributor". This strategy was used for all the indicators that may have similar correlation problems.
- The number of times that pairs of article contributors have interacted in any article talk page, divided by group size.

4.3.2.4 Group heterogeneity

In this study, group heterogeneity will be operationalized in accordance to the TIP theory as follows:

- Tenure heterogeneity: The standard deviation of the tenure of the article contributors. Member tenure is defined as the elapsed amount of time since a member's first contribution in the English Wikipedia.

- Member contribution size heterogeneity (group production): The standard deviation of the average contribution size of the article contributors (from the sample) to which the members have contributed.

- Member contribution frequency heterogeneity (group production): The standard deviation of the frequency of contribution per article contributor.

- Member comment size heterogeneity (group well-being): The standard deviation of the average talk page contribution size per article contributor (from all the article talk pages of the sample to which the members have contributed). In this study, "comment" refers to a contribution to an article talk page, as distinct from a contribution to the article itself.

- Member comment frequency heterogeneity (group well-being): The standard deviation of the frequency of talk page contributions per article contributor.

- Member user comment size heterogeneity (member support): The standard deviation of the average contribution size to user pages and user talk pages per article contributor (from all the user pages and user comment pages of the users included in the sample).

- Member user comment frequency heterogeneity (member support): The standard deviation of the frequency of user page and user talk page contributions per article contributor (from all the user pages and user comment pages of the users included in the sample).

4.3.2.5 Member competency

Member competency will be measured by the following items:

- Proportion of Administrators: The number of individuals that have administrator status in Wikipedia that are present in an article, divided by the total group size. Administrators are wikipedians who have special rights in the web-based encyclopedia. They can edit Wikipedia's home page and other protected pages; they can permanently delete; they have better tools for reverting a page to an earlier version; they have the authority to enforce rulings by the Wikipedia Arbitration Committee; they can block or unblock user IP addresses from access to Wikipedia; and they have some other administrative rights. It is important to pinpoint that before any administrator status request, a user must have sufficiently contributed to the Wikipedia community so as to be recognized by the other users who will then agree on the promotion. As a consequence, once elected, an administrator is a person who has been implicated in the virtual community life for a while, and who has acquired some strong knowledge about the Wikipedia article writing process.

- Average member tenure: The number of days since a contributor's first contribution to an article or its corresponding talk page, in the Wikipedia community.

- Peer Recognition: It is the custom for wikipedians to recognize each other for hard work and helpfulness by awarding each other "barnstars". To give the award to someone, a wikipedian simply places one of the barnstar images on the recipient's user talk page, and cites why it was awarded. Peer recognition will be measured by

the total number of barnstars received by contributors to an article, divided by group size.

4.3.2.6 Member activeness

We measured member activity in line with the TIP theory group activity processes. It consists of describing a wikipedian's frequency of activity during their entire membership in the Wikipedia community in terms of group production, group well-being and member support functions. The chosen operationalization is the following:

- Average number of article contributions per day (group production).
- Average size of article contributions per day (group production).
- Average number of comments per day (group well-being).
- Average size of comments per day (group well-being).
- Average number of comments in user pages or user talk pages per day (member support).
- Average size of comments in user pages or user talk pages per day (member support).

4.3.3 Process variables

The following sections will introduce the operationalization of the three group process variables: group production, group well-being and member support.

4.3.3.1 Group production

Group production is directly related to the article production function, which is basically the posting of contributions to the article page. Group production function will be measured according to the following items:

- Equality of participation: This is the extent to which contributors have participated equally in the writing of an article. It will be measured by the standard deviation of total contribution size per user per article.
- Level of effort: This is how much effort contributors have put into writing an article, measured by:
  - total article size in number of characters
  - number of individual contributions
  - average contribution size by each contributor
- Production Rate: This describes the production process through time in terms of duration, speed and regularity of production. We will consider the following measures:
  - Production duration: Age of the article (in days)
  - Production rate (number): Average number of contributions per day
  - Production rate (size): Average size of contributions per day
  - Production regularity: Standard deviation of the number of contributions per month.

- Emergent leadership: This describes the emergence of a particular user who appears to lead the production process of an article. It will be measured in terms of:
  - Emergent leader's degree of participation: The number of contributions of the user who has contributed the most (in number of contributions) divided by the total number of contributions by all contributors of that article.
  - Emergent leader's size of participation: The total size of the contributions of the user who have contributed the most (in total size of contributions) divided by the total size of the contributions by all contributors of that article.

4.3.3.2 Group well-being

As group members interact often in terms of exchanging information, coordinating tasks, exchanging points of view—that is to say, to increase group consciousness through the article production process—article quality may increase significantly as group processes become increasingly efficient. To measure such group processes, several items will be considered:

- Information exchange: This is the amount of information that group members have exchanged during their interaction in the talk page. Two measures will be used:
  - Overall size of the talk page
  - Average comment size in talk page

- Degree of interaction: This is the extent to which all group members have participated in the talk page in order to exchange ideas, viewpoints and recommendations. Several measures will be used:
  - The percentage of group members (article contributors) who have participated in the article talk page.
  - Total number of contributors who have written comments in the article talk page.
  - Average number of comments per contributor in the article talk page.
- Level of conflict: This refers to the extent to which an article has raised issues related to conflicts and disputes that have occurred through the production of an article. Level of conflict will be measured by the number of times an article has been flagged as incorporating a conflict. Three cases may occur: the neutrality of an article is contested[8], the accuracy of an article is disputed[9], and a request for external comment is required[10] due to interpersonal conflicts between wikipedians. For each case, Wikipedia keeps track of each dispute in designated pages. Level of conflict will be measured by the number of times an article has appeared in one of those pages.
- Coordination effort: In order to coordinate their efforts, contributors may use a feature called a "to-do" list, which lists improvements that are suggested for the article. This list is maintained by editors, writers, reviewers or readers as a way to focus collaborative efforts. As such, they represent a tentative consensus, helping

[8] http://en.wikipedia.org/wiki/Category:NPOV_disputes
[9] http://en.wikipedia.org/wiki/Wikipedia:Accuracy_dispute
[10] http://en.wikipedia.org/wiki/Wikipedia:Requests_for_comment

improve the efficiency of the editing process. Coordination effort will be measured by the number of tasks that have been added to the to-do list of an article.

- Emergent Facilitator: This refers to the presence of a contributor who has managed and facilitated members' interaction in the talk page. An emergent facilitator differs from an emergent leader in that the facilitator focuses on managing the interactions among group members, whereas the emergent leader is strictly related to the production aspect of articles.

- Two measures will be included:

  - The maximum number of comments posted by any one contributor in the article talk page divided by the total number of posted comments.

  - The maximum total comment size posted by any one contributor in the article talk page divided by the total size of posted comments.

4.3.3.3 Member support

As contributors participate in the writing and editing of an article, long-term collaboration may give birth to the building of relationships among contributors. Indeed, relationship building is a major aspect of the member support function in the sense that such relationships are seen as individual gains and payoffs of the article production process. Thus, member support will be operationalized as interpersonal communications not directly related to the creation of specific articles. This notion refers to the extent to which wikipedians who have worked on the same articles exchange personal information and communicate through their personal user page and user talk page. Because user pages

and user talk pages are modified independently of the articles writing process, it was impossible to track the member support activities directly associated to a specific article. As a consequence, member support had some inherent measurement constraints and limitations in the context of Wikipedia. The following measures will be used:

- The number of comments that a contributor has received on his or her user page and user talk page from the other contributors to the considered article, divided by group size.
- The total size of comments that a contributor has received on his or her user page and user talk page from the other contributors to the considered article, divided by group size.

4.3.4 Reflective and formative constructs

The measure of theoretical constructs can be performed with two different approaches. First, a classical view that is mostly used is to assume that the variation in the scores on measures of a construct is a function of the true score of the construct plus an error term. Thus, the underlying latent construct, said to be a *reflective construct*, causes the observed variation in the measures, assuming a direction of causality from the latent variable to its measures (Jarvis et al., 2003; Nunnally, 1978). However, this approach is conceptually inappropriate for certain latent variables, defined as *formative constructs*, for which the direction of causality should rather be viewed as emanating from the measures to the corresponding construct (Jarvis et al., 2003). Formative constructs, also called composite latent variables (Blalock, 1964), are measures that form or cause the

creation or change in a latent variable (Chin, 1998). An example of formative construct may be the beliefs construct of the theory of reasoned action (Fishbein & Ajzen, 1975) in which each individual belief causes the overall belief construct (Jarvis et al., 2003).

Both types of latent variables were used to test the theoretical model. Shared Experience, Member Competency and Member Support were defined as reflective variables. They will be measured with items that are expected to covary, thus justifying the use of a reflective approach. Furthermore, Group Size and Effectiveness will be further treated as reflective constructs for they are operationalized as single-indicator constructs.

Organization Support, Group Heterogeneity, Group Activeness, Group Production and Group Well-being were conceptualized as formative constructs. The choice of the formative constructs has several explanations. First, the Time, Interaction and Performance theory (McGrath, 1991) introduces three distinct function levels that are group production, group well-being and member support. For each of them, McGrath provides a set of different and independent activities. Section 3.2.2 has introduced the activities that have been investigated in group research in the MIS field. Following those theoretical considerations, it is crucial to pinpoint that for each group process function, those activities are not affected by an underlying construct. Instead, these activities create and cause change in their corresponding group process latent construct. As a consequence, Group Production and Group Well-being were defined and operationalized as formative constructs. Furthermore, Group Activeness and Group Heterogeneity also rely on the three group function levels and were thus conceptualized and operationalized as formative constructs. Organizational Support is also seen as the combination of the

different means an organization supports a group work. This latent construct was thus also conceptualized and operationalized as a formative variable.

In defining the measurement items, a particular effort was done in regards to the formative constructs to define as many indicators as possible. Indeed, to assess formative constructs, it is important to measure such constructs with a large number of indicators to adequately tap into all the dimensions of the constructs (Bollen & Lennox, 1991). The nature of the constructs in summarized in Table 2.

Table 2: Constructs Operationalization and Measurement

Variable Category	Variable Name	Operationalization	Measure Name	Measurement
INPUT	*Organizational Support* (formative)	Collaboration of the week/ Article Improvement Drive/ Requests for Expansion/ Pages Needing Attention/ Cleanup Requests	in_orgsup_sum	The total number of times an article has received a "collaboration of the week" support and per or "improvement drive" and per or "page needing attention" and per or "request for expansion" request and per or "cleanup" requests.
		Peer Review	in_orgsup_peerrev	The total number of times an article has received a "peer review" request.
	Group Size	Open Group Size	in_size_open	The total number of contributors of an article.
	Shared Experience (reflective)	Article Shared Experience	in_shexp_art	The number of times that pairs of contributors of the article has interacted in the **article** itself divided by group size.
		Discussion Shared Experience	in_shexp_talk	The number of times that pairs of contributors of the article has interacted in the **article talk page** divided by group size.
	Group Heterogeneity (formative)	Tenure Heterogeneity	in_tenurstd	The standard deviation of Member Tenure.
		Group Production Heterogeneity	in_membprodstd_contsize	Member Contribution Size Heterogeneity: The standard deviation among group members of the average **size** of contributions per user.
			in_membprodstd_contfreq	Member Contribution Frequency Heterogeneity: The standard deviation among group members of the average **number** of contributions per day per user.
		Group Well-Being Heterogeneity	in_gpwellstd_comsize	Member Comment Size Heterogeneity: the standard deviation among group members of the average **size** of comments per user in article talk pages.
			in_gpwellstd_comfreq	Member Comment Frequency Heterogeneity: the standard deviation among group members of the average **number** of comments per day per user in article talk pages.
		Member Support Heterogeneity	in_membsupstd_comsize	Member User Comment Size Heterogeneity: the standard deviation among group members of the average **size** of comments per user in other users' user pages and user talk pages.
			in_membsupstd_comfreq	Member User Comment Frequency Heterogeneity: the standard deviation among group members of the average **number** of comments per day per user in other users' user pages and user talk pages.
	Member competency (reflective)	Average Member Tenure	in_tenuravg	The number of days since a contributor's first contribution in the Wikipedia community.
		Presence of Administrators	in_admin	The number of administrators that have contributed of the article divided by the total number of contributors of the article.
		Peer Recognition	in_peerrec_barn	The total number of barnstars attributed to the contributors of the article divided by the total number of contributors of the article.
			in_peerrec_awd	The total number of awards (wikithanks) attributed to the contributors of the article divided by the total number of contributors of the article.
	Member Activeness (formative)	Group Production Activeness	in_membprodavg_contsize	The average among group members of the average **size** of contributions per user.
			in_membprodavg_contfreq	The average among group members of the average **number** of contributions per day per user.
		Group Well-Being Activeness	in_gpwellavg_comsize	The average among group members of the average **size** of comments per user in article talk pages.
			in_gpwellavg_comfreq	The average among group members of the average **number** of comments per day per user in talk pages.
		Member Support Activeness	in_membsupavg_comsize	The average among group members of the average **size** of comments per user in other users' user pages and user talk pages.
			in_membsupavg_comfreq	The average among group members of the average **number** of comments per day per user in other users' user pages and user talk pages.

Table 2: Constructs Operationalization and Measurement (continued)

Variable Category	Variable Name	Operationalization	Measure Name	Measurement
PROCESS	Group Production (formative)	Participation Equality	pr_partic_contsizestd	The standard deviation of the total **size** of contributions per contributor of the article
			pr_partic_contnbstd	The standard deviation of the **number** of contributions per contributors of the article.
		Effort Level	pr_effort_totsize	Total size of the article in number of characters.
			pr_effort_contnb	The total number of contributions to the article.
			pr_effort_contavgsize	The average size of the contributions to the article by all contributors.
		Production Rate	pr_prodpat_artage	Production process length: The amount of time, in days, that has passed from the date of the first contribution.
			pr_prodpat_sizerate	Average size of contributions per day.
			pr_prodpat_rate	Production Speed: Average number of contributions per day.
			pr_prodpat_regul	Production Regularity: Standard deviation of the number of contributions per month.
		Emergent Leader	pr_lead_deg	Emergent Leader's Degree of Paiticipation: The **% number** of edits of the user who has posted more contributions than any other contributor.
			pr_lead_size	Emergent Leader's Size of Participation: The **% size** of total contributions of the user who has posted more contributions than any other contributor.
	Group Well-Being (formative)	Information Exchange	pr_infoex_talksize	Length of the talk page.
			pr_infoex_comavgsize	Average size of a comment in a talk page.
		Interaction Degree	pr_interac_avgcomusr	The % group members (contributors of the article) who have participated in the talk page.
			pr_interac_membperc	The average number of comments per user.
			pr_interac_comnb	The number of comments in the talk page.
		Coordination	pr_coord	The number of "to-do" tasks found in the entire history of the article talk page.
		Level of Conflict	pr_confl_sum	The number of times an article has been quoted in "NPOV", "request for comments page", and other pages related to conflicting issues.
		Emergent Facilitator	pr_facil_deg	Emergent Facilitator's Degree of Participation: The **% number** of comments of the user who has posted more comments than any other contributor.
			pr_facil_size	Emergent Facilitator's Size of Participation: The **% size** of total comments of the user who has posted more comments than any other contributor.
	Member Support (reflective)	Relationship Building	pr_relbuild_nb	The **number** of comments that a contributor has received on his per her user page and user talk page from the other contributors of the considered article divided by the total number of contributors of the article.
			pr_relbuild_size	The total **size** of the comments that a contributor has received on his per her user page and user talk page from the other contributors of the considered article divided by the total number of contributors of the article.
OUTPUT	Effectiveness	Article Quality	feat_nom_score	Article quality level among: regular article with no nomination, featured article nominees that were not accepted, and featured articles.

5 *DATA ANALYSIS AND RESULTS*

This chapter presents the data analysis procedure that was used and also describes the results provided by the statistical analysis. First, descriptive statistics will give an overall picture of the measurement items used to assess the model. Second, the statistical analysis choice will be presented and justified. Then, the detailed steps of the instrument validation will be described. Finally, the theoretical model will be tested in accordance to the selected statistical approach, thus identifying the group input and group process factors that lead to high-quality projects.

5.1 Descriptive statistics

Table 3 reports the minimum, maximum, mean, median, mode and standard deviation of each measurement item. Several descriptive statistics give an overall picture of Wikipedia articles. For instance, the average article length in the chosen sample was found to be 12,653 characters (σ=12,298). It has to be kept in mind that the articles that were eligible for our sample, needed to have a minimum size of at least 5,000 characters. Furthermore, some groups were assigned a size of zero when no registered user participated in the corresponding article production processes.

Table 3: Descriptive Statistics

	MIN	MAX	Mean	Median	Mode	Std Dev
Organizational Support						
in_orgsup_sum	0	2	0.05	0	0	0.231
in_orgsup_peerrev	0	1	0.03	0	0	0.175
Group Size						
in_size_open	0	1055	27.56	15	4	40.589
Shared Experience						
in_shexp_art	0	86534	3289.10	1064.5	0	5770.746
in_shexp_talk	0	2515	37.55	5	0	102.177
Group Heterogeneity						
in_tenurstd	0	610	277.60	303.501	0	104.855
in_membprodstd_contsize	0	61069	602.13	453.396	0	1127.239
in_membprodstd_contfreq	0	10	1.36	1.297915	0	0.816
in_gpwellstd_comsize	0	13667	724.86	415.961	0	996.100
in_gpwellstd_comfreq	0	3	0.19	0.173926	0	0.139
in_membsupstd_comsize	0	108264	1688.58	343.829	0	8366.491
in_membsupstd_comfreq	0	9	0.93	0.836767	0	0.631
Member Competency						
in_tenuravg	0	1338	515.76	531.857	0	151.658
in_admin	0	230	9.29	5	1	12.869
in_peerrec_barn	0	156	6.31	3	1	8.791
in_peerrec_awd	0	65	2.50	1	0	3.765
Member Activeness						
in_membprodavg_contsize	0	18184	474.87	399.6575	0	439.209
in_membprodavg_contfreq	0	12	1.19	1.0644	0	0.674
in_gpwellavg_comsize	0	28507	513.29	439.922	0	512.820
in_gpwellavg_comfreq	0	2	0.13	0.116944	0	0.096
in_membsupavg_comsize	0	78092	825.68	432.479	0	3163.451
in_membsupavg_comfreq	0	10	0.74	0.658575	0	0.487

Moreover, the sample was characterized by a mean article age of 652 days (σ=412 days), an average of 70 contributions (σ=148 contributions) whereas the average number of comments in talkpages was 14 (σ=75 comments). On average, the articles have 28 contributors (σ=41).

Table 3: Descriptive Statistics (continued)

	MIN	*MAX*	*Mean*	*Median*	*Mode*	*Std Dev*
Group Production						
pr_partic_contsizestd	0	155181	2575.70	1528.67	0	4812.672
pr_partic_contnbstd	0	134	3.78	2.03522	0	5.753
pr_effort_totsize	5000	455809	12653.36	8609	5166	12298.613
pr_effort_contnb	0	6151	70.44	33	11	148.284
pr_effort_contavgsize	0	76831	563.51	257.772	0	1582.673
pr_prodpat_artage	1	1536	652.39	608	1151	412.195
pr_prodpat_sizerate	3	22122	78.34	17.83495	11	535.674
pr_prodpat_rate	0	30	0.15	0.064568	0	0.633
pr_prodpat_regul	0	2420	7.18	3.37065	0	28.528
pr_lead_deg	0	1	0.36	0.288462	0.5	0.242
pr_lead_size	0	164	0.79	0.73833	1	1.901
Group Well-Being						
pr_infoex_talksize	0	2800559	6628.44	261.5	0	46103.749
pr_infoex_comavgsize	0	14767	229.53	115	0	473.736
pr_interac_avgcomusr	0	1	0.10	0.057143	0	0.136
pr_interac_membperc	0	25	0.28	0.066667	0	0.738
pr_interac_comnb	0	3366	13.88	1	0	75.208
pr_coord	0	1	0.01	0	0	0.099
pr_confl_sum	0	3	0.03	0	0	0.178
pr_facil_deg	0	1	0.31	0.222222	0	0.351
pr_facil_size	0	184	0.40	0.234223	0	2.195
Member Support						
pr_relbuild_nb	0	58216	391.08	55	0	1476.997
pr_relbuild_size	0	16777215	226898.62	22839	0	828093.000
Effectiveness						
feat_nom_score	0	2	0.17	0	0	0.504

5.2 Statistical analysis technique

This section explains the statistical methods that will be employed to analyze the data to test the hypotheses of this study.

The literature review has allowed us to develop a group model that relies on group input, process and outputs that can be tested using partial least squares (PLS) analysis.

This statistical technique appears to be particularly appropriate for it permits to test multi-level models that feature interactions and multiple-cause effects.

In addition, because of the exploratory character of this research project and because the proposed model has not been tested in the literature, PLS appeared as an appropriate statistical tool (Zhu & Kraemer, 2005). Equally importantly, PLS can handle both reflective and formative constructs (Chin, 1998) which makes it the most appropriate statistical tool for this research project. Indeed, it was shown that several attempts to explicitly model formative indicators in an SEM analysis (which could otherwise have been a possible alternative), lead to identification problems (Chin, 1998).

5.3 Assessment of the measurement model (outer model)

In order to validate the measurement model, a distinction had to be made between reflective constructs and formative ones. Bollen and Lennox (1991) pointed out that the traditionally used methods for assessing both construct reliability and validity are not appropriate for formative constructs, where the direction of causality is posited to flow from the measures to the constructs (Diamantopoulos & Winklhofer, 2001; Jarvis et al., 2003). Following the procedures used in prior analysis using PLS (Bagozzi, 1994; Carmines & Zeller, 1979; Chin, 1998; Diamantopoulos & Winklhofer, 2001; Gefen & Straub, 2005; Hulland, 1999), the adequacy of the reflective constructs was assessed through the following tests: item reliability, construct reliability, convergent validity and discriminant validity. The formative constructs were validated by looking at item collinearity and discriminant validity (Hulland, 1999).

PLS Graph 3.00 was used to run the overall model. In PLS, both reliability and validity tests of a measurement model are assessed through the use of confirmatory factor analysis (CFA) (Cronbach, 1951). Basically, CFA consists of conducting a factor analysis of all items in the instrument where each item loads on its corresponding construct. CFA generates a factor-analytic loading and a weight (beta regression coefficient) for each measurement item in relation to the latent variable they address. The results of the CFA simulation are presented in Table 4 and Table 5.

5.3.1 Reflective latent variables

The following sections describe the different steps of the validation of the reflective latent variables of the measurement instrument.

5.3.1.1 Item reliability

The purpose of an item reliability test is to verify whether each of the measurement items acts consistently as a measure of the corresponding construct.

In PLS, individual item reliability for reflective latent variables is assessed by examining the loadings of the measures with their respective construct (provided by CFA). A common rule of thumb that has been employed by many researchers is to accept items with loadings of 0.7 or more (Chin, 1998). This criterion corresponds to a level of shared variance between a construct and its measure that is greater than the corresponding error variance (Carmines & Zeller, 1979). Since loadings are correlations, this implies that more than 50 percent of the variance in the observed variable (meaning

the square of the loading) is due to the construct (Hulland, 1999). Among the reflective constructs, two items of the Member Competency latent variable had loadings below the 0.7 limit and were thus dropped (in_peerrec_barn and in_peerrec_awd). The other loadings were all satisfactory, as shown in Table 4. The p-values were found in deriving them from the corresponding t-values obtained after having run a bootstrap procedure with 200 resamples.

Table 4: Confirmatory Factor Analysis (CFA) results for the reflective constructs

CONSTRUCT	ITEM	INITIAL INSTRUMENT		REFINED INSTRUMENT		
		Weights	Loadings	Weights	Loadings	p-value
Shared Experience	in_shexp_art	0.502	**0.867**	0.539	**0.882**	0.000
	in_shexp_talk	0.618	**0.914**	0.583	**0.900**	0.000
Member Competency	in_tenuravg	0.598	**0.840**	0.780	**0.958*** **	0.000
	in_admin	0.356	**0.808**	0.335	**0.755*** **	0.000
	in_peerrec_barn	0.200	**0.441**			
	in_peerrec_awd	0.246	**0.496**			
Member Support	pr_relbuild_nb	0.566	**0.957**	0.566	**0.961*** **	0.000
	pr_relbuild_size	0.487	**0.941**	0.487	**0.945*** **	0.000

*: 0.05 significance level
**: 0.01 significance level
***: 0.001 significance level

5.3.1.2 Construct reliability

Construct reliability is defined as the assurance that the items posited to measure a construct are sufficiently related to be reliable when considered as a set of items (Cronbach, 1951).

Table 5: Confirmatory Factor Analysis (CFA) results for the formative constructs

CONSTRUCT	ITEM	INITIAL INSTRUMENT		REFINED INSTRUMENT		
		Weights	Loadings	Weights	Loadings	p-value
Org. Support	in_orgsup_sum	0.157	0.203	0.145***	0.190	0.000
	in_orgsup_peerrev	0.980	0.988	0.983***	0.990	0.000
Group Heterogeneity	in_tenurstd	0.468	0.681	0.607***	0.778	0.000
	in_membprodstd_contsize	0.280	0.391	0.233***	0.360	0.000
	in_membprodstd_contfreq	-0.028	0.310	0.026	0.339	0.122
	in_gpwellstd_comsize	0.251	0.433	0.197***	0.387	0.000
	in_gpwellstd_comfreq	0.469	0.728	0.375***	0.670	0.000
	in_membsupstd_comsize	0.104	0.126	0.075***	0.104	0.000
	in_membsupstd_comfreq	0.196	0.601	0.176***	0.561	0.000
Member Activeness	In_membprodavg_contsize	-0.007	0.025	-0.086*	-0.046	0.017
	in_membprodavg_contfreq	-0.579	-0.124	-0.597***	-0.155	0.000
	in_gpwellavg_comsize	0.171	0.283	0.179**	0.286	0.003
	in_gpwellavg_comfreq	1.009	0.823	1.039***	0.803	0.000
	in_membsupavg_comsize	0.165	0.112	0.129***	0.076	0.000
	in_membsupavg_comfreq	0.065	0.484	0.005	0.428	0.471
Group Production	pr_partic_contsizestd	-0.118	0.301			
	pr_partic_contnbstd	-0.042	0.296	0.236***	0.332	0.000
	pr_effort_totsize	0.247	0.525	0.373***	0.549	0.000
	pr_effort_contnb	0.807	0.934			
	pr_effort_contavgsize	-0.030	-0.081	-0.083***	-0.155	0.000
	pr_prodpat_artage	0.235	0.573	0.563***	0.771	0.000
	pr_prodpat_sizerate	-0.022	-0.072	-0.063**	-0.134	0.005
	pr_prodpat_rate	0.028	0.243	0.070***	0.201	0.000
	pr_prodpat_regul	-0.066	0.406	0.150*	0.408	0.039
	pr_lead_deg	-0.068	-0.390	-0.274***	-0.549	0.000
	pr_lead_size	0.097	0.198	0.144**	0.253	0.004
Group Well-Being	pr_infoex_talksize	-0.140	0.765	0.735***	0.860	0.000
	pr_infoex_comavgsize	0.196	0.388	0.178***	0.458	0.000
	pr_interac_avgcomusr	0.386	0.428	0.404***	0.543	0.000
	pr_interac_membperc	-0.414	0.417	-0.162	0.520	0.087
	pr_interac_comnb	1.079	0.909			
	pr_coord	0.151	0.264	0.229***	0.306	0.000
	pr_confl_sum	0.050	0.438	0.145**	0.498	0.003
	pr_facil_deg	-0.105	0.034			
	pr_facil_size	0.005	0.048	0.002	0.088	0.446

In general, construct reliability has been measured by researchers using internal consistency (Cronbach's alpha coefficient). The Cronbach's alpha should be above 0.60 in exploratory research and above 0.70 for confirmatory (Gefen & Straub, 2005). In PLS,

construct reliability is more appropriately assessed with two other indicators. First, composite reliability (or Rho) proposed by Fornell and Larcker (1981) is not influenced by the number of items in the scale as is Cronbach's alpha but only by the relative loadings of the items. The minimum composite reliability level is 0.8. As shown inTable 6, all the constructs had a composite reliability above the 0.8 limit.

Table 6: Composite reliability

	Reliability (ρ)
Shared Experience	0.886
Member Competency	0.851
Member Support	0.948

Second, another pertinent measure of construct validity is the average variance extracted (AVE) which is computed as the average of the squared loadings of each item in a construct measure. AVE measures how well a theoretical latent variable explains the variation in the set of items that are used to measure it. Fornell and Larcker (1981) posited that the AVE of a construct should be at least 0.5 for an acceptable measure of a latent variable. As seen in Table 7, all the AVEs of the reflective constructs scored much higher than the minimum of 0.5 which confirmed a high reliability of each construct.

Table 7: Correlation between latent constructs (square root of AVE in the leading diagonal)

	AVE	*Group Size*	*Shared Exp*	*Member Comp*	*Member Sup*	*Effectiv*
Group Size	**1.000**	**1.000**				
Shared Exp	**0.795**	0.517	**0.892**			
Member Comp	**0.743**	0.129	0.458	**0.862**		
Member Sup	**0.900**	0.492	0.557	0.281	**0.949**	
Effectiv	**1.000**	0.422	0.434	0.050	0.303	**1.000**

5.3.1.3 Convergent validity

In PLS, both convergent and discriminant validity are assessed through factorial validity. Gefen and Straub (2005) proposed a practical guide to test construct validity using PLS in the case of reflective constructs; those guidelines were rigorously followed to test construct validity in our measurement model. Convergent validity is shown when each measurement item correlates strongly with its assumed theoretical construct (Gefen & Straub, 2005). To assess convergent validity, each of the measurement items loading should have a significant t-value. Typically, the p-value of this t-value should be significant at least at a 0.05 level. All the loadings were found to have a p-value lower than 0.001, thus confirming high construct validity (as seen in Table 4).

5.3.1.4 Discriminant validity

In contrast to convergent validity, discriminant validity is shown when each measurement item correlates weakly with all other constructs except for the one to which it is theoretically associated. Discriminant validity was tested according to the two methods recommended by Gefen and Straub (2005).

The first step is to assess if any item highly scores on their theoretically assigned factor and not highly on the others. The results of the cross-loading analysis are presented in Table 8. All the concerned items scored much higher in their corresponding latent variables.

Moreover, establishing discriminant validity also requires an appropriate AVE analysis which consists of testing whether the square root of every AVE is higher than the correlation coefficients (Pearson's factor) of any pair of the latent constructs. As shown in Table 7, the square root AVE of all reflective latent constructs was consistently higher, thus confirming a high level of discriminant validity.

Thus, as a result of the assessment of the reflective variables of the measurement model, two items were dropped because of item reliability issues (in_peerrec_barn and in_peerrec_awd), leading to fully valid and reliable measures of the reflective constructs.

Table 8: Cross-loadings

	REFLECTIVE VARIABLES					FORMATIVE VARIABLES				
	Group Size	*Shared Exp*	*Memb. Comp*	*Memb. Sup*	*Effect.*	*Org Sup*	*Group Heter.*	*Memb. Activ*	*Group Prod*	*Group Well*
in_size_open	**1.0000**								0.6930	
in_shexp_art		**0.8824**							0.6332	
in_shexp_talk		**0.9003**							0.5148	
in_tenuravg		0.4689	**0.9592**							
in_admin		0.2768	**0.7521**							
pr_relbuild_nb		0.5720		**0.9568**						
pr_relbuild_size		0.4780		**0.9411**						
feat_nom_score		0.4336			**1.0000**					
in_orgsup_sum						**0.1900**				0.0995
in_orgsup_peerrev						**0.9896**				0.2908
in_tenurstd			0.5811				**0.7898**			
in_membprodstd_contsize							**0.3424**		0.2062	
in_membprodstd_contfreq		0.3196					**0.3443**			
in_gpwellstd_comsize							**0.3793**		0.2178	
in_gpwellstd_comfreq							**0.6687**	0.6866		
in_membsupstd_comsize				0.1330			**0.1047**			
in_membsupstd_comfreq							**0.5610**	0.4219		
in_membprodavg_contsize		-0.0612	-0.0864				0.0575	**-0.0543**	-0.0912	
in_membprodavg_contfreq	-0.1545							**-0.1490**		
in_gpwellavg_comsize							0.2120	**0.2758**		
in_gpwellavg_comfreq							0.4327	**0.8128**		
in_membsupavg_comsize				0.1093				**0.0742**		
in_membsupavg_comfreq							0.2457	**0.4349**		
pr_partic_contnbstd					0.3203				**0.3299**	
pr_effort_totsize	0.4138								**0.5470**	
pr_effort_contavgsize							-0.1795		**-0.1605**	
pr_prodpat_artage		0.6254							**0.7899**	
pr_prodpat_sizerate			-0.1479				-0.1616		**-0.1382**	
pr_prodpat_rate	0.2200								**0.1895**	
pr_prodpat_regul	0.3794								**0.3569**	
pr_lead_deg							-0.4798		**-0.5629**	
pr_lead_size	0.1722								**0.2157**	
pr_infoex_talksize	0.5181									**0.8587**
pr_infoex_comavgsize									0.2856	**0.4685**
pr_interac_avgcomusr								0.3193		**0.5565**
pr_interac_membperc								0.2487		**0.5292**
pr_coord					0.1568					**0.3145**
pr_confl_sum	0.2812									**0.5133**
pr_facil_size									0.0477	**0.0662**

When the highest correlation factor of an indicator is inside its corresponding construct, the second highest correlation factor is displayed. In case the highest correlation factor of an indicator is not inside its corresponding construct, then all the correlation factors that are also higher are displayed.

5.3.2 Formative latent variables

As pointed out by Loch et al. (2003), there is little guidance in the literature on detailed validation procedures of formative constructs other than Diamantopoulos and Winkofler's (2001) suggestions. A delicate issue is that the items of formative constructs are supposed to cover the whole scope of a latent variable which constrains instrument refinement. However, it does not mean that no item purification is possible by excluding items (Diamantopoulos & Winklhofer, 2001). It simply means that at the indicator specification stage, the chosen indicators should be sufficiently inclusive to fully capture the domain of content of a construct (Diamantopoulos & Winklhofer, 2001).

When unobservable, underlying constructs are seen as giving rise to associated measures, it is appropriate to talk about item reliability and convergent validity. In the case of formative latent variables, it is not appropriate (Bollen & Lennox, 1991; Hulland, 1999). In fact, formative indicators of the same latent variable may have positive, negative or even zero correlation (Bollen & Lennox, 1991). As a consequence, reliability is not a meaningful concept when applied to formative constructs (Cohen et al., 1990).

Furthermore, as the nature of formative constructs renders internal consistency inappropriate, "the best we can do ... is to examine how well the index relates to measures of other variables" (Bagozzi, 1994, p. 333). Following both Diamantopoulos and Winkofler's (2001) guidelines and Bagozzi's suggestions (Bagozzi, 1994), item collinearity and discriminant validity were examined.

5.3.2.1 Item collinearity

Because the formative measurement model is based on multiple regression, the stability of the indicator coefficients (or beta weights) might be strongly affected by a strong inter-correlation between the items of a same construct, called collinearity, which must be carefully assessed when using PLS (Gefen et al., 2000). For each formative construct, the Variance Inflation Factors (VIF) were computed. Variance Inflation Factors measure the impact of collinearity among the independent variables in a regression model on the precision of estimation. It expresses the degree to which collinearity among the predictors degrades the precision of an estimate. The literature has used several common cut-off points ranging from 10 to 2.5. It was decided that all the items whose VIF value were above 2.5 had to be removed from the measurement instrument. Two items were thus dropped: pr_partic_contsizestd of Group Production, and pr_interac_comnb of Group Well-being.

5.3.2.2 Discriminant validity

Finally, discriminant validity was assessed by using a cross-loading analysis. One item was dropped due to an overly high loading in another construct than the one it belongs to (pr_effort_contnb with a loading of 0.87 with Group Size). Another item was also dropped because it scored much higher in another construct than in its own (pr_facil_deg scoring 0.20 in Group Heterogeneity whereas scoring only 0.034 in its own construct: Group Well-being). Moreover, the analysis revealed that some other items had higher loadings in other constructs than in their own one. From a theoretical perspective, elimination of indicators carries the risk of changing a formative construct itself, by

removing important aspects of it (Diamantopoulos & Winklhofer, 2001). The items that showed an acceptable amount of loading in other constructs were finally kept in the final instrument, the largest difference between the correlation factors never exceeding 0.04. Those items were: in_gpwellstd_comfreq (own construct correlation: 0.669, maximum correlation in other construct: 0.687), in_membsupstd_comsize (own: 0.105, max. other: 0.133), in_membprodavg_contsize (own: -0.054, max. other: -0.091), in_membprodavg_contfreq (own: -0.149, max. other: -0.154), in_membsupavg_comsize (own: 0.074, max. other: 0.109), pr_effort_contavgsize (own: -0.161, max. other: -0.180), pr_prodpat_sizerate (own: -0.138, max. other: -0.162), pr_prodpat_rate (own: 0.190, max. other: 0.220), and pr_prodpat_regul (own: 0.357, max. other: 0.379).

As a conclusion of the assessment of the formative constructs of the measurement model, out of the 34 formative indicators, 2 indicators were dropped for multi-collinearity reasons (pr_partic_contsizestd and pr_interac_comnb) and two more were dropped because of discriminant validity issues (pr_effort_contnb and pr_facil_deg), resulting in refined formative construct measures and a fully validated measurement instrument.

5.4 Assessment of the structural model (inner model)

PLS Graph 3.00 was used in order to assess the structural model. Statistical significance was assessed using a bootstrap procedure with 200 resamples. The structural model test consists of estimating the path coefficients, which indicate the strength of the relationship between independent and dependent variables to test the hypotheses stated in Table 1. The squared multiple correlation (R^2) for each endogenous construct in the

theoretical model corresponds to the amount of variance explained by independent variables. These values were interpreted similarly to the R^2 provided by regression model (Chwelos et al., 2001; Wixom & Watson, 2001). Because PLS does not generate an overall goodness-of-fit index, the validity of a model is assessed by examining R^2 and the structural paths (Chwelos et al., 2001; Wixom & Watson, 2001).

All the exogenous variables explained 65.6% of the variation in Group Production, 34.7% of the variation of Group Well-being, and 38.4% of the variation of Member Support. Overall, 29.0% of the variation of the primary dependent variable, Effectiveness, was explained by the variables of the model. The significance of the R^2 coefficients was determined using the test described by Falk and Miller (1992, p. 72):

$$F = \frac{R^2 / m}{(1-R^2)(N-m-1)}$$

N is the number of cases and m the number of items that measure the construct under study. All R^2 were found to be significant at the 0.001 level.

Assuming a significance level of 0.05 and according to past studies using path analysis, it was decided to adopt the following categorization of the results. An hypothesis was said to be "not supported" in two different cases: either the path was not significant or its strength did not exceed 0.1 even though highly significant. A hypothesis was said to be "weakly supported" when the corresponding path was at least 0.1 but less than 0.2 and significant. Moreover, a hypothesis was claimed to be "supported" when its path was at least 0.2 and less than 0.3 and significant, following Chin's recommendations (1998). Finally, a hypothesis was qualified as "strongly supported" when its

corresponding path was greater than 0.3 and significant. The assessment of the significance of the structural paths was determined by a bootstrap procedure with 200 resamples (Table 9). A graph summarizing the findings is presented in Figure 3.

The findings support several hypotheses of the model. Out of the 27 hypothesized paths, 19 were found highly significant ($p \leq 0.001$), 1 very significant ($p \leq 0.01$), 1 significant ($p \leq 0.05$) and 6 were not significant. The standardized paths of statistically significant relationships range from −0.142 to 0.536 with several paths exceeding the suggested minimum of 0.2 (Chin, 1998). However, even though structural paths with a coefficient lower than 0.2 are not as strong, past research has considered paths at least above 0.1 as meaningful and very insightful (when statistically significant) (Croteau & Bergeron, 2001; Plouffe et al., 2001; Ridings et al., 2002).

The exploratory nature of this research project makes it even more important for each of these paths that are found highly significant may unveil an unexplored open content research avenue that may trigger further research efforts, even though the standard paths might only be as low as 0.1.

Organizational Support was found to be positively related to Group Well-being and Effectiveness, thus partially supporting H_{1a}, and fully supporting H_{1b}. The corresponding paths were 0.146 with Group Well-being and 0.126 with Effectiveness (path= 0.126).

Table 9: Path coefficients and R^2 for overall model

Predictor Construct	*Predicted Construct*	*Path*	*t-value*	*p-value*	*Hypothesis*
Organizational Support	→ Group Production	0.022	1.443	0.075	$H_{1a(i)}$
	→ Group Well-Being	**0.146*** **	5.831	0.000	**$H_{1a(ii)}$**
	→ Member Support	0.045***	3.734	0.000	$H_{1a(iii)}$
	→ Effectiveness	**0.126*** **	7.351	0.000	H_{1b}
Group Size	→ Group Production	**0.463*** **	11.844	0.000	**$H_{2a(i)}$**
	→ Group Well-Being	**0.536*** **	5.857	0.000	**$H_{2a(ii)}$**
	→ Member Support	**0.280*** **	16.135	0.000	**$H_{2a(iii)}$**
	→ Effectiveness	**0.177*** **	5.845	0.000	**H_{2b}**
Shared Experience	→ Group Production	**0.216*** **	11.908	0.000	**$H_{3a(i)}$**
	→ Group Well-Being	-0.072	1.469	0.071	$H_{3a(ii)}$
	→ Member Support	**0.369*** **	21.559	0.000	**$H_{3a(iii)}$**
	→ Effectiveness	**0.360*** **	15.801	0.000	**H_{3b}**
Group Heterogeneity	→ Group Production	**0.316*** **	11.936	0.000	**$H_{4a(i)}$**
	→ Group Well-Being	-0.017	0.751	0.226	$H_{4a(ii)}$
	→ Member Support	-0.033**	2.643	0.004	$H_{4a(iii)}$
	→ Effectiveness	-0.073***	4.927	0.000	H_{4b}
Member Competency	→ Group Production	0.055***	5.747	0.000	$H_{5a(i)}$
	→ Group Well-Being	-0.031	1.380	0.084	$H_{5a(ii)}$
	→ Member Support	0.077***	8.101	0.000	$H_{5a(iii)}$
	→ Effectiveness	-0.142***	15.990	0.000	H_{5b}
Member Activeness	→ Group Production	-0.095***	5.707	0.000	$H_{6a(i)}$
	→ Group Well-Being	**0.153*** **	5.428	0.000	**$H_{6a(ii)}$**
	→ Member Support	**0.109*** **	6.053	0.000	**$H_{6a(iii)}$**
	→ Effectiveness	**0.117*** **	9.323	0.000	**H_{6b}**
Group Production	→ Effectiveness	0.075*	2.252	0.012	H_7
Group Well-Being	→ Effectiveness	0.014	0.441	0.330	H_8
Member Support	→ Effectiveness	-0.002	0.139	0.445	H_9

(In bold: significant paths ($p \leq 0.05$) whose strength exceeds or equals 0.1).

	R^2	*F-value*	*p-value*
Group Production	0.6560***	2116.744	0.000
Group Well-Being	0.3468***	757.857	0.000
Member Support	0.3837***	3111.998	0.000
Effectiveness	0.2898***	4079.725	0.000

***: 0.05 significance level**
****: 0.01 significance level**
*****: 0.001 significance level**

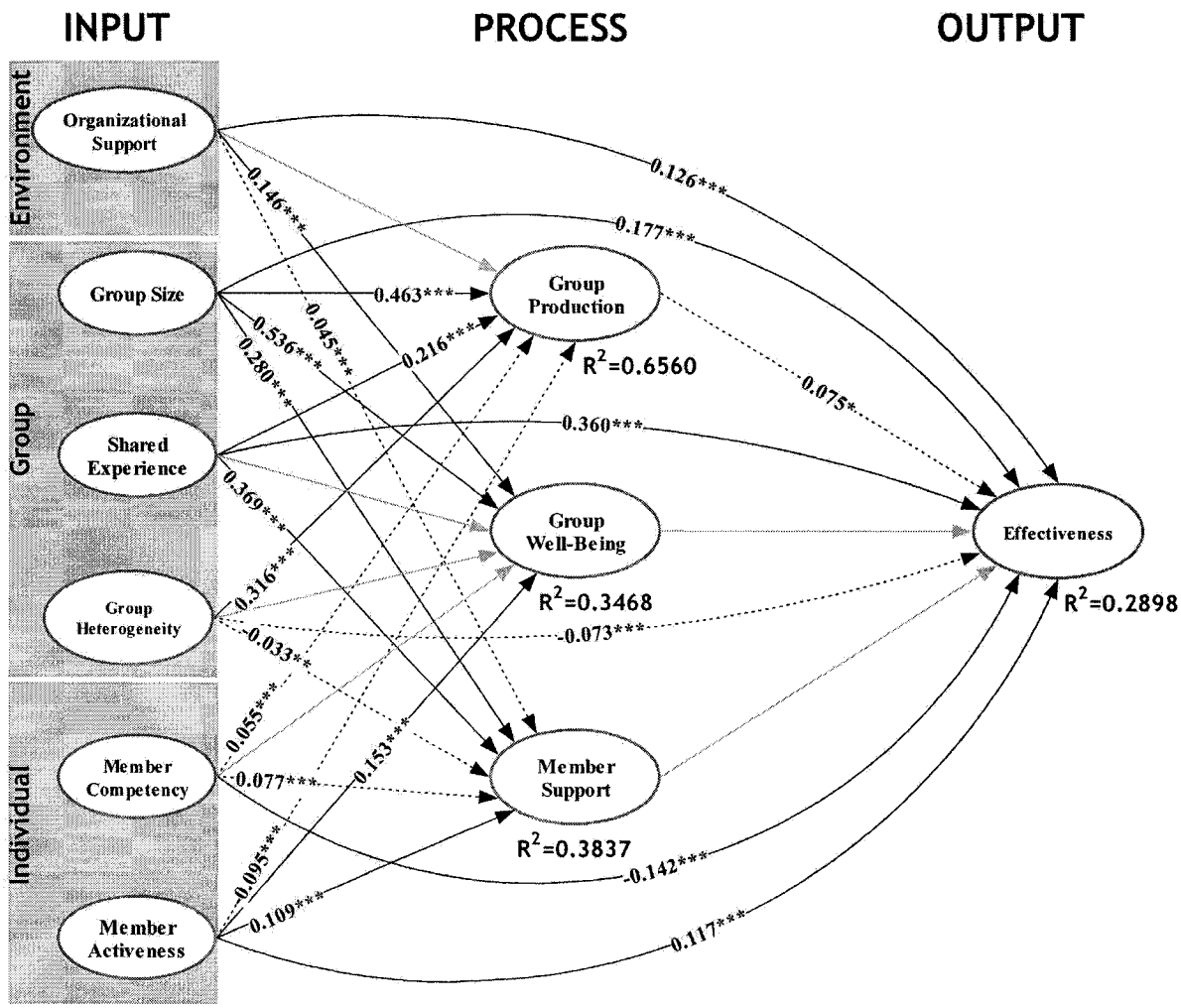

Figure 3: Verified paths in model of group processes in open content communities

Group Size was found to be among the most explanatory and significant exogenous variables, being positively related to all endogenous variables. H_{2a} and H_{2b} were completely supported: the path with Group Production was 0.463, the one with Group Well-being: 0.536, Member Support: 0.280, and Effectiveness: 0.177.

Shared Experience also provided highly explanatory results. It had a highly significant positive relationship with Group Production (with a path of 0.216), Member

Support (path 0.369), thus partially supporting H_{3a}, and effectiveness with a path of 0.360 (H_{3b} strongly supported).

The results found with Group Heterogeneity partially support H_{4a} for it had a strong and significant positive relationship with Group Production (path = 0.316).

Member Competency provided unexpected results. It was found not be related to Group Production, Group Well-being and Member Support, contradicting H_{5a}. However, a negative and significant path was found between Member Competency and Effectiveness (-0.142), which finding will be further discussed in the following chapter.

The path analysis of Member Activeness yielded partially hypothesized results. It was found to have a positive and significant relationship with both Group Well-being and Member Support (with paths of respectively 0.153 and 0.109) but a negative and significant path was found with Group Production (-0.095). H_{6b} was weakly supported for member activeness, with a path of 0.117 for Effectiveness.

Finally, neither H_7 nor H_8 were supported, leading to questions about the explanatory power of the process variables on open content group effectiveness. The hypothesis results are summarized in Table 10.

Table 10: Hypothesis Results

Hyp.	*Content*	*Results*
H_{1a}	Organizational support is positively related to the group process variables: (i) group production; (ii) group well-being; and (iii) member support.	(i) and (iii) not supported **(ii) weakly supported**
H_{1b}	Organizational support is positively related to effectiveness.	**weakly supported**
H_{2a}	Group size is positively related to the group process variables: (i) group production; (ii) group well-being; and (iii) member support.	**(i, ii) strongly supported (iii) supported**
H_{2b}	Group size is positively related to effectiveness.	weakly supported
H_{3a}	Shared experience is positively related to the group process variables: (i) group production; (ii) group well-being; and (iii) member support.	**(i) supported (iii) strongly supported**
H_{3b}	Shared Experience is positively related to effectiveness.	**strongly supported**
H_{4a}	Group heterogeneity is positively related to the group process variables: (i) group production; (ii) group well-being; and (iii) member support.	**(i) strongly supported** (ii) and (iii) not supported
H_{4b}	Group heterogeneity is positively related to effectiveness.	not supported
H_{5a}	Member competency is positively related to the group process variables: (i) group production; (ii) group well-being; and (iii) member support.	not supported
H_{5b}	Member competency is positively related to effectiveness.	not supported (inverse effect)
H_{6a}	Member activeness is positively related to the group process variables: (i) group production; (ii) group well-being; and (iii) member support.	**(ii) and (iii) weakly supported** (i) inverse effect
H_{6b}	Member activeness is positively related effectiveness.	**weakly supported**
H_7	Group production is positively related to effectiveness.	not supported
H_8	Group well-being is positively related to effectiveness.	not supported
H_9	Member support is positively related to effectiveness.	not supported

6 *DISCUSSION AND CONCLUSION*

This chapter discusses the main findings of this study based on the group input and group process variables that were found highly explanatory in relation to OC group effectiveness. Then, the implications and further research avenues are presented and will shed light on future research on open content (OC) communities and groups. The main limitations of this project will be briefly discussed, followed by an overall conclusion.

6.1 Main findings

This research project has revealed several valuable discoveries about open content group behaviour and overall effectiveness. This section introduces and discusses them.

6.1.1 Predictive ability of open content group inputs

Organizational support was found to be positively related to group well-being (path of 0.146, $p < 0.001$). In Wikipedia, organizational support is mainly expressed through peer review requests which means that people are encouraged to join the current group of an article in order to discuss how it could be modified and improved. However, group production was not found to be affected by organizational support. As a consequence, it seems that organizational support in the context of open content groups plays a role in stimulating the social activities of group tasks by increasing discussion, coordination and exchange of information but not the production task itself. Furthermore,

the member support function was not found to be affected by organizational support which indicates that members are influenced by organizational support in group activities that are only directly related to the group task.

Moreover, the results confirmed that organizational support plays a role on open content group effectiveness. This confirms similar previous findings of the management literature which found that organizational support leads to an increase in employees' creativity (Scott & Bruce, 1994). As a conclusion, organizational support in the context of open content groups seems to be more about affecting quality rather than quantity of group tasks.

The influence of group size, as expected, is essential in open content projects. A larger pool of contributors is related to more group production, group well-being and member support. Group size has a synergistic effect on open content group activities. Open content group effectiveness is also significantly influenced by group size, thus confirming past studies that showed that larger groups work better (Dennis & Wixom, 2001). In the case of Electronic Meeting Systems (EMS), larger groups generated more unique and more high-quality ideas, while members were more satisfied when using the EMS (Gallupe et al., 1992).

The strongest predictor of both group processes and effectiveness was found to be shared experience. First, a high degree of shared experience stimulates open groups to produce more. Nonetheless, shared experience was not found to be related to group well-being function. Since people with a high level of shared experience, are used to work together, there is less need for interaction and synchronization for people know each other, and know what they have to do. Furthermore, as people have contributed to articles

together, they tend to know each other more, thus increasing friendly interaction through relationship building processes (member support function). As a consequence, the synergy among the group members with high shared experience help them in focusing mainly on their production tasks, without having to interact much in relation to the group task but leaving some space and time for personal friendly exchanges.

Group heterogeneity also provided some interesting results. However, it is important to note that Group Heterogeneity was operationalized in terms of actual behaviours, rather than in terms of innate qualities or competencies of group members. Group production was found to be positively related to group heterogeneity, which confirms past findings in the literature. Indeed, it has been shown that such membership diversity expands the available resources to the group, therefore increasing the likelihood of improved productivity (Shaw, 1981). Group heterogeneity was not found to have any influence on the group well-being and member support functions. A possible explanation is that a high degree of group heterogeneity among group members might create a certain lack of connectivity, leading to fewer exchanges among members. Thus, as users of such groups do not share the same interests in the virtual community, there is no motivation for interaction. In addition, group heterogeneity was not found to be directly related to effectiveness as opposed to findings in traditional group research (Aladwani et al., 2000; Paul et al., 2004; Steiner, 1972). A possible explanation is that the size of open content groups is much larger than the one of regular work groups (the average group size was 27.6 people with a standard deviation of 40.6). As a consequence, open content groups tend to be heterogeneous in nature due to group size effects. In contrast, past group research that investigated group heterogeneity used much smaller group sizes (less than

10 people) where the amount of heterogeneity was restricted so that the experiments could be more easily controlled.

Member Competency was not found to be related to any of the group process variables. Operationalized in terms of average member tenure and average member status in the community (administrators or regular contributors), a possible explanation is that in virtual communities, competency and experience are two related notions. An administrator is a user who has been an active Wikipedia contributor for a while and who is generally a known and trusted member of the community. Such a person tends to be responsible for dealing with high level tasks such as discussing and voting policies, solving conflicts, electing and voting featured article candidates or any suggestion or recommendation by any member of the community.

That a significant negative relationship was found between member competency and effectiveness is quite surprising, and warrants some further analysis. First, a cross-item correlation analysis revealed that administrators are more likely to be present in articles that have had conflicts (with a correlation factor of 0.25). Similar findings were found by past group research. Indeed, Kelly & Bostrom (1998) showed that experienced group members are more inclined to deal with socio-emotional issues. Depending on the level and type of conflict, two different consequences may rise. First, a minor conflict attracts the attention of some administrators that contribute to solve it, which could substantially increase the quality of the article. The second possibility is the case of articles with an inherently conflictual nature due to its topic (political, religious, ideological, etc.). Administrators also strive to resolve such a type of conflict but those issues may never be resolved, so that such an article might never attain the required

quality criteria to reach featured status. As a result, the presence of a high proportion of administrators does not lead to any quality achievements. Nevertheless, it is important to insist that the proportion of administrators is only a single measurement item of the Member Competency construct. A correlation test revealed that the presence of a large number of administrators in an article, taken alone, overall increases the quality (with a high correlation of 0.450). As a result, the presence of administrators is not negative in itself, but further research is needed to clarify this issue.

6.1.2 Group functions in open content groups

65.6% of the variation of group production was explained by the input variables. Group Size was the most influential factor (path = 0.463) followed by Group Heterogeneity (0.316) and then Shared Experience (0.216). This indicates that open content group production is stimulated by larger groups, a high activity diversity of its members and also a high degree of shared experience. Group production was not found to be influenced by organizational support, member competency, or member activeness.

Around 35% of the variation of group well-being was explained. The most influential factors were Group Size (0.536), Member Activeness (0.153), and Organizational Support (0.146). This indicates that the open content group well-being function is affected by larger groups, a high degree of activity of group members, and support from the open content community. Group well-being was not found to be influenced by shared experience, group heterogeneity, or member competency.

38.4% of the variation of the member support function was explained by the proposed input factors. The most influential factors were found to be: Shared Experience (0.369), Group Size (0.280), and Member Activeness (0.109). This indicates that in the context of open content groups, relationship building is mainly encouraged and increased in groups with a high degree of shared experience, in larger groups, and also in groups which members are more active in the community. Member Support was not found to be related to organizational support, group heterogeneity, or member competency.

6.1.3 Predicting open content group effectiveness

Open content group effectiveness was found to be influenced by all of the three categories of group inputs: environmental, group and individual. An important result was revealed in regards to the role played by both group size and shared experience. Indeed, even though significant support was found in regards to the overall effectiveness of larger groups, their members have to know each other through past collaboration. In other words, group size matters, but more importantly group members must have shared some previous experience in order to be effective when working together. Moreover, open content group effectiveness is related to the presence of active members, who seem to be the harder workers of the community. The support of the community is also important in increasing the quality of open content projects as it stimulates group members to collaborate and interact more.

However, no direct relation was found with any of the group processes, the effects of the input variables fully explaining the variation in open content group effectiveness. However, a supplementary PLS analysis of the process variables as independent variables

and Group Effectiveness as the dependent variable, revealed that around 25% of the variation of Effectiveness was explained, the largest path being Group Production (0.368), then Group Well-being (0.135) and then Member Support (0.131). First, this confirms the basic belief that the main factor for producing quality open content group projects is through the group production function, which basically means doing the actual job. Second, this analysis fully justifies the explanatory and behavioural importance of the process variables even though it clearly shows that no extra variation of open content group effectiveness is explained by these process variables beyond the variation explained by the input variables.

6.2 Implications and further research

This study has brought a strong empirical confirmation of the importance of group size in relation to open content group effectiveness. More importantly, some new light has been shed in regards to the important role of shared experience. As pointed out in the literature review, neither traditional group nor open content research has established clear results about its importance. Further research in both contexts is thus needed in order to confirm the essentialness of shared experience in explaining group effectiveness.

In addition, this research project has investigated some classical group factors such as organizational support, member heterogeneity, and member activeness. Those variables provided some significant findings that had never been shown in the literature. Such issues had never been addressed in the open content context either, and should be further investigated.

This research project is among the first to propose an operationalization of the Time, Interaction and Performance theory in the MIS field (Dennis & Reinicke, 2004, being the first such study). Moreover, it was the first attempt to conceptualize both group production and group well-being functions as formative variables, providing a broader and more encompassing view of those concepts as opposed to previous studies such as Dennis and Reinicke for instance, who captured only one aspect of both group well-being and member support (group production was not included in their study) by using reflective measures, thus focusing on the correlation of the measurement items. As a consequence, more research is needed in order to have a clearer picture of the TIP group process functions, through a precise operationalization of the dimensions of the constructs. This research project introduced an operationalization that explained around 66% of the group production construct, 35% of the group well-being one and 38% of member support which leaves room for investigation of other explanatory factors of these constructs.

Member activeness was also found to be a good explanatory factor of open content group behaviour. Due to the shortage of prior theoretical support, this issue should be further investigated in the context of open content communities, as this study found that it played a significant role on both group processes and outcomes. This notion may be particularly insightful and may shed some light in the broader context of virtual communities in which group works differ from traditional group tasks in professional environments.

Other measures of group performance could have been used in this research project. Even though it was decided that effectiveness through article featured status was

the most appropriate and accurate measure, efficiency measures may also provide insightful and complementary results. For example, a suggested efficiency measure in the context of Wikipedia may be the amount of time before reaching featured status nomination.

Overall, this research project has attempted to provide general results about both group behaviour and outcomes of open content groups. Wikipedia, an open-content web-based encyclopedia was used as a sample for this study, but there is now a need for validating and extending the results to other open content groups in general, such as open source software development.

6.3 Limitations

This research project had some limitations that need to be noted. First, the operationalization of some of the constructs such as group heterogeneity and member competency had to be restricted to actual member activities within the context of the studied community, as demographic variables such as knowledge, skills or educational background were not available. As a consequence, the results must be interpreted in light of the specific operationalization employed, and should not be assumed to be applicable with significantly different operationalizations.

Another limiting aspect concerns the measurement of the member support function. It was not technically possible to evaluate the member support activities that were directly linked to a specific article. As a consequence, the results provided regarding

member support have to be carefully assessed. Promising insights were found but further research is needed.

In addition, due to computational and memory constraints, only a sample of 10,000 out of more than 525,000 articles were used in order to validate and test the model. In addition, only the articles with a minimum size of 5000 characters were taken into consideration. The results may have slightly differed by including all the articles, or by using some other sampling strategy.

6.4 Conclusion

This study investigated the group variables (input and process) that lead to high-quality products in open content communities. Based on the Input-Process-Output perspective and on the Time, Interaction, and Performance Theory proposed by McGrath (1991), a theoretical model was developed. The model assessed the relationship between input variables (organizational support, group size, shared experience, group heterogeneity, member competency, and member activeness), process variables (group production, group well-being and member support) and group effectiveness. A sample of 10,000 group tasks in Wikipedia, an open content encyclopedia, was used in a quantitative study. This study demonstrated the overall synergistic effect of group size and shared experience on both group process variables and group effectiveness. This project also showed (1) the positive effect of group heterogeneity on group production; (2) the positive effect of organizational support and member activeness on group well-being; (3) the positive effect of member activeness on member support; and finally (4)

the positive effect of organizational support, and member activeness on group effectiveness.

This research project has shed further light on how open content projects create quality products. A significant contribution has been brought to both MIS research and group research by providing group insights from social psychology in investigating the factors that lead to high quality information-based products. The advent of the Internet has been challenging the time and geographical constraints of group collaboration by enabling new practices that rely on new streams of thoughts. Open content is a nascent phenomenon which has started delivering on some of its attractive promises.

BIBLIOGRAPHY

Adkins, M., Burgoon, M., & Jr., J. F. N. (2003). Using group support systems for strategic planning with the United States Air Force. *Decision Support Systems, 34*(3), pp. 315-337.

Aiken, M., Krosp, J., Shirani, A., & Martin, J. (1994). Electronic brainstorming in small and large groups. *Information and Management, 27*, pp. 141–149.

Aladwani, A. M., Rai, A., & Arkalgud, R. (2000). Formal participation and performance of the system development group: The role of group heterogeneity and group-based rewards. *Database for Advances in Information Systems, 31*(4), pp. 25-41.

Alonzo, M., & Aiken, M. (2004). Flaming in electronic communication. *Decision Support Systems, 36*(3), pp. 205-213.

Anakwe, U. P., Anandarajan, M., & Igbaria, M. (1999). Information technology usage dynamics in Nigeria: An empirical study. *Journal of Global Information Management, 7*(2), pp. 13-22.

Arrow, H., & McGrath, J. E. (1995). Membership dynamics in groups at work: A theoretical framework. In B. M. S. L. L. Cummings (Ed.), *Research in Organizational Behavior,* (Vol. 17, pp. 373-411). Greenwich, CT: JAI Press.

Bagozzi, R. P. (1994). *Advanced Methods in Marketing Research* Cambridge, Mass.: Blackwell.

Baldi, S., Heier, H., & Mehler-Bicher, A. (2003). Technical opinion: Open courseware and open source software. *Communications of the ACM, 46*(9), pp. 105-107.

Barkhi, R., Jacob, V. S., & Pirkul, H. (1999). An Experimental Analysis of Face to Face versus Computer Mediated Communication Channels. *Group Decision and Negotiation, 8*, pp. 325-347.

Barkhi, R., Jacob, V. S., & Pirkul, H. (2004). The influence of communication mode and incentive structure on GDSS process and outcomes. *Decision Support Systems, 37*(2), pp. 287-305.

Batenburg, R. S., & Bongers, F. J. (2001). The role of GSS in participatory policy analysis. A field experiment. *Information & Management, 39*(1), pp. 15 - 30.

Bergquist, M., & Ljungberg, J. (2001). The power of gifts: organizing social relationships in open source communities. *Information Systems Journal, 11*(4), pp. 305-320.

Blalock, H. M. (1964). *Causal Inferences in Nonexperimental Research* Chapel Hill, NC: University of North Carolina Press.

Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110*(2), pp. 305–314.

Bose, U., & Paradice, D. B. (1999). The Effects of Integrating Cognitive Feedback and Multi-attribute Utility-Based Multicriteria Decision-Making Methods in GDSS. *Group Decision and Negotiation, 8*, pp. 157-182.

Burke, K., & Chidambaram, L. (1994). *Development in Electronically-Supported Groups: A Preliminary Longitudinal Study of Distributed and Face-to-face Meetings.* Paper presented at the Twenty-seventh Hawaii International Conference on System Sciences (HICSS).

Burke, K., & Chidambaram, L. (1999). How much bandwidth is enough? A longitudinal examination of media characteristics and group outcomes. *MIS Quarterly, 23*(4), pp. 557-579.

Carillo, K. D., & Okoli, C. (2005). Open Source Software Communities. In S. Dasgupta (Ed.), *Encyclopedia of Virtual Communities and Technologies*. George Washington University, USA: Idea Group Reference.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and Validity Assessment: Quantitative Applications in Social Science* Beverly Hills, California: Sage Publications.

Cedergren, M. (2003). Open content and value creation. *First Monday, 8*(8), pp.

Chengalur-Smith, S., & Sidorova, A. (2003, December 14 - 17, 2003). *Survival of Open-Source Projects: A Population Ecology Perspective.* Paper presented at the 24th International Conference on Information Systems 2003, Seattle, Washington.

Chidambaram, L. (1996). Relational Development in Computer-Supported Groups. *MIS Quarterly, 20*(2), pp. 143-176.

Chin, W. W. (1998). Issues and opinion on structural equation modeling. *MIS Quarterly, 22*(1), pp. vii–xvi.

Chwelos, P., Benbasat, I., & Dexter, A. S. (2001). Research Report: Empirical Test of an EDI Adoption Model. *Information Systems Research, 12*(3), pp. 304-321.

Cohen, P., Cohen, J., Teresi, J., Marchi, M., & Velez, C. N. (1990). Problems in the measurement of latent variables in structural equations causal models. *Applied Psychological Measurement, 14*, pp. 183-196.

Cronbach, L. J. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika, 16*, pp. 297-334.

Croteau, A.-M., & Bergeron, F. (2001). An Information Technology Trilogy: Business Strategy, Technological Deployment and Organizational Performance. *Journal of Strategic Information, 2*, pp. 77-99.

Dennis, A. R., & Garfield, M. J. (2003). The adoption and use of GSS in project teams: Toward more participative processes and outcomes. *MIS Quarterly, 27*(2), pp. 289.

Dennis, A. R., George, J. F., Jessup, L. M., Nunamaker, J. J. F., & Vogel, D. R. (1988). Information Technology to Support Electronic Meetings. *MIS Quarterly, 12*(4), pp. 591-624.

Dennis, A. R., Jay, E. A., William, G. H., & Edward, D. W. (1999). Structuring time and task in electronic brainstorming. *MIS Quarterly, 23*(1), pp. 95-108.

Dennis, A. R., & Reinicke, B. A. (2004). Beta versus VHS and the acceptance of electronic brainstorming technology. *MIS Quarterly, 28*(1), pp. 1-20.

Dennis, A. R., & Wixom, B. H. (2001). Investigating the moderators of the group support systems use with meta-analysis. *Journal of Management Information Systems, 18*(3), pp. 235.

DeSanctis, G., & Gallupe, R. B. (1987). A Foundation for the Study of Group Decision Support Systems. *Management Science, 33*(5), pp. 589-600.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research, 38*, pp. 269-277.

Diker, V. G., & Scholl, H. J. (2001, January 3-6, 2001). *The art of leveraging: how powerful nonlinear feedback processes can restructure rapidly growing technology and knowledge industries.* Paper presented at the 34th Hawaii International Conference on System Sciences, Island of Maui, Hawaii.

Easton, G., Easton, A., & Belch, M. (2003). An experimental investigation of electronic focus groups. *Information & Management, 40*(8), pp. 717-727.

Edwards, J. (1998). The changing face of freeware. *IEEE Computer, 31*(10), pp. 11 - 13.

Falk, R. F., & Miller, N. B. (1992). *A Primer for Soft Modeling* Akron, OH: University of Akron Press.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research* Reading, MA: Addison-Wesley.

Fitzgerald, B. (2004). A critical look at open source. *IEEE Computer, 37*(7), pp. 92 - 94.

Fjermestad, J., & Hiltz, S. R. (1999). An assessment of group support systems experiment research: Methodology and results. *Journal of Management Information Systems, 15*(3), pp. 7- 15.

Fornell, C., & Larcker, D. (1981). Structural Equation Models With Unobserved Variables and Measurement Error. *Journal of Marketing Research, 18*, pp. 39-50.

Gallupe, R. B., Dennis, A. R., Cooper, W. H., Valacich, J. S., Bastianutti, L. M., & Jay F. Nunamaker, J. (1992). Electronic Brainstorming and Group Size. *Academy of Management Journal, 35*(2), pp. 350-369.

Gallupe, R. B., DeSanctis, G., & Dickson, G. W. (1988). Computer-Based Support For Group Problem-Finding: An Experimental Investigation. *MIS Quarterly, 12*(2), pp. 277-297.

Gefen, D., & Straub, D. W. (2005). A Practical Guide to Factorial Validity Using PLS-Graph: Tutorial and Annotated Example. *Communications of the Association for Information Systems, 16*(5), pp. 91-109.

Gefen, D., Straub, D. W., & Boudreau, M. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the AIS, 4*(7), pp. 1–76.

Ghosh, R. A., Glott, R., Krieger, B., & Robles, G. (2002). *Free/Libre and Open Source Software: Survey and Study.* Brussels: Report of the FLOSS Workshop on Advancing the Research Agenda on Free/Open Source Software, European Commission.

Goodman, P. S., Ravlin, E. C., & Argote, L. (1986). *Current thinking about groups: Setting the stage for new ideas.*Designing effective work groups, San Francisco: Jossey-Bass.

Hackman, J. R. (1987). The design of work teams. In J. W. Lorsch (Ed.), *Handbook of Organizational Behavior* (pp. 315-342). Englewood Cliffs, NJ: Prentice- Hall.

Hackman, J. R., & Morris, C. G. (1975). Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 8, pp. 45-99). NewYork: Academic Press.

Hann, I.-H., Roberts, J., Slaughter, S. A., & Fielding, R. (2002, December 15 - 18, 2002). *Economic Incentives for Participating in Open Source Software Projects.* Paper presented at the 23rd International Conference on Information Systems, Barcelona, Spain.

Hars, A., & Ou, S. (2001, January 3-6, 2001). *Working for free? Motivations of participating in open source projects.* Paper presented at the 34th Hawaii International Conference on System Sciences, Island of Maui, Hawaii.

Hiltz, S. R., Dufner, D. K., Holmes, M. E., & Poole, S. M. (1991). Distributed Group Support Systems: Social Dynamics and Design Dilemmas. *Journal of Organizational Computing, 2*(1), pp. 135-159.

Holsapple, C. W., & Luo, W. (1999). Effects of Experience-based Work Patterns in a GSS Environment. *Group Decision and Negotiation, 8*, pp. 305-324.

Huang, W. W., & Wei, K. K. (1997). Task as a moderator for the effects of group support systems on group influence processes. *European Journal of Information Systems, 6*(4), pp. 208-217.

Huang, W. W., & Wei, K. K. (2000). An empirical investigation of the effects of Group Support Systems (GSS) and task type on group interactions from an influence perspective. *Journal of Management Information Systems, 17*(2), pp. 181-207.

Huang, W. W., Wei, K. K., & Tan, B. C. Y. (1999). Compensating effects of GSS on group performance. *Information & Management, 35*(4), pp. 195.

Huang, W. W., Wei, K. K., Watson, R. T., & Tan, B. C. Y. (2003). Supporting virtual team-building with a GSS: An empirical investigation. *Decision Support Systems, 34*(6), pp. 359.

Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: a review of four recent studies. *Strategic Management Journal, 20*(2), pp. 195 - 204.

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research. *Journal of Consumer Research, 30*, pp. 199-218.

Jelassi, M. T., & Beauclair, R. A. (1987). An Integrated Framework for Group Decision Support Systems Design. *Information and Management, 13*, pp. 143-153.

Kayworth, T. R., & Leidner, D. E. (2001). Leadership effectiveness in global virtual teams. *Journal of Management Information Systems, 18*, pp. 7-41.

Kelly, G. G., & Bostrom, R. P. (1998). A facilitator's general model for managing socioemotional issues in group support systems meeting environments. *Journal of Management Information Systems, 14*(3), pp. 23-44.

Kelly, J. R. (1988). Entrainment in group interaction and task performance. In. In J. E. McGrath (Ed.), *The social psychology of time: New perspectives* (pp. 89-110). Thousand Oaks, CA: Sage Publications.

Kelly, J. R., & McGrath, J. E. (1985). Effects of time limits and task types on task performance and interaction in four-person groups. *Journal of Personality and Social Psychology, 49*, pp. 395-407.

Kim, Y., Hiltz, S. R., & Turoff, M. (2002). Coordination Structures and System Restrictiveness in Distributed Group Support Systems. *Group Decision and Negotiation, 11*, pp. 379-404.

Kraut, R. E. (2003). Applying social psychological theory to the problems of group work. In J. M. Carroll (Ed.), *Hci Models, Theories, & Frameworks: Toward a Multidisciplinary Science.*

Kutsko, J., & Smith, J. Y. (1991). *Effectiveness Measures for Distributed Teams Using Electronic Meeting Technology: The Larson/Lafasto Instrument.* Paper presented at the the Twenty-Fourth Hawaii International Conference on System Sciences.

Kwok, R. C. W., & Khalifa, M. (1998). Effect of GSS on knowledge acquisition. *Information & Management, 34*(6), pp. 307-315.

Lakhani, K., & von Hippel, E. (2003). How Open Source Software Works: 'Free,' User-to-User Assistance. *Research Policy, 32*(6), pp. 923-943.

Lakhani, K., Wolf, B., Bates, J., & DiBona, C. (2002). *The Boston Consulting Group Hacker Survey*. Boston: Boston Consulting Group and Open Source Developers Network.

LaMastro, V. (2001). Influence of perceived institutional and faculty support on college students' attitudes and behavioral intentions. *Psychological Reports, 88*(2), pp. 567-580.

Lee, F. S. L., Vogel, D., & Limayem, M. (2003). Virtual community informatics: A review and research agenda. *JITTA: Journal of Information Technology Theory and Application, 5*(1), pp. 47-61.

Lerner, J., & Tirole, J. (2002). Some simple economics of open source. *Journal of Industrial Economics, 50*(2), pp. 197-234.

Limayem, M., & DeSanctis, G. (2000). Providing decisional guidance for multicriteria decision making in groups. *Information Systems Research, 11*(4), pp. 386–401.

Ljungberg, J. (2000). Open source movements as a model for organising. *European Journal of Information Systems, 9*(4), pp. 208-216.

Loch, K. D., D.W., S., & Kamel, S. (2003). Diffusing the Internet in the Arab world: The role of social norms and technological culturation. *IEEE Transactions on Engineering Management, 50*(1), pp. 45-63.

Lurey, J. S., & Raisinghani, M. S. (2001). An empirical study of best practices in virtual teams. *Information & Management, 38*, pp. 523-544.

Time Magazine (2005, November 09, 2005), available at.

Malhotra, A., Majchrzak, A., Carman, R., & Lott, V. (2001). Radical innovation without collocation: A case study at Boeing-Rocketdyne. *MIS Quarterly, 25*, pp. 229.

Massey, A., Montoya-Weiss, M., & Hung, Y.-T. (2002, January 07 - 10, 2002). *Synchronizing Pace in Asynchronous Global Virtual Project Teams.* Paper presented at the 35th Annual Hawaii International Conference on System Sciences, Big Island, Hawaii.

Mathiyalakan, S. (2002). A methodology for controlled empirical investigation of membership continuity and change in GDSS groups. *Decision Support Systems, 32*(3), pp. 279-296.

McGrath, J. E. (1984). *Groups: Interaction and performance* Englewood Cliffs: Prentice-Hall.

McGrath, J. E. (1991). Time, Interaction, and Performance (TIP). A Theory of Groups. *Small Group Research, 22*, pp. 147-174.

McGrath, J. E., Arrow, H., Gruenfeld, D. H., Hollingshead, A. B., & O'Connor, K. M. (1993). Group Tasks and Technology: The Effects of Experience and Change. *Small Group Research, 24*(3), pp. 406-420.

McHaney, R., & Cronan, T. P. (2000). Toward an empirical understanding of computer simulation implementation success. *Information & Management, 3*, pp. 135-151.

Montoya-Weiss, M. M., Massey, A. P., & Song, M. (2001). Getting it together: Temporal coordination and conflict management in global virtual teams. *Academy of Management Journal, 44*(6), pp. 1251-1262.

Neus, A. (2001). *Managing information quality in virtual communities of practice: Lessons learned from a decade's experience with exploding Internet communication.* Paper presented at the 6th International Conference on Information Quality at MIT, Boston, Massachusetts.

Niederman, F., & Volkema, R. J. (1999). The Effects of Facilitator Characteristics on Meeting Preparation, Set Up, and Implementation. *Small Group Research, 30*(3), pp. 330-360.

Nunamaker, J. F., Dennis, A. R., Valacich, J. S., Vogel, D., & George, J. F. (1991). Electronic meeting systems. *Communications of the ACM, 34*(7), pp. 40 - 61.

Nunnally, J. C. (1978). *Psychometric theory (2nd Edition)* New York: McGraw-Hill.

O'Reilly, T. (1999). Lessons from open-source software development. *Communications of the ACM, 42*(4), pp. 32-37.

Oetzel, J. G. (2001). Self-construals, communication processes, and group outcomes in homogeneous and heterogeneous groups. *Small Group Research, 32*(1), pp. 19–54.

Parent, M., Gallupe, R. B., Salisbury, W. D., & Handelman, J. M. (2000). Knowledge creation in focus groups: Can group technologies help? *Information & Management, 38*(1), pp. 47-58.

Paul, S., Priya, S., Imad, S., & Peter, P. M. (2004). Impact of heterogeneity and collaborative conflict management style on the performance of synchronous global virtual teams. *Information & Management, 41*(3), pp. 303-321.

Pauleen, D. J. (2003). An Inductively Derived Model of Leader-Initiated Relationship Building with Virtual Team Members. *Journal of Management Information Systems, 20*(3), pp. 227-256.

Perens, B. (1997). The Open Source Definition. 2005, from http://www.opensource.org/docs/definition.php

Pinsonneault, A., & Kraemer, K. L. (1990). The Effects of Electronic Meetings on Group Processes and Outcomes: An Assessment of the Empirical Research. *European Journal of Operational Research, 46*(2), pp. 143-161.

Plouffe, C. R., Hulland, J., & Vandenbosch, M. (2001). Research Report: Richness Versus Parsimony in Modeling Technology Adoption Decisions –Understanding Merchant Adoption of a Smart Card-Based Payment System. *Information Systems Research, 12*(2), pp. 208-222.

Pollard, C. (2003). Exploring Continued and Discontinued Use of IT: A Case Study of OptionFinder, a Group Support System. *Group Decision and Negotiation, 12*(3), pp. 171-193.

Poole, M. S., & DeSanctis, G. (1990). Understanding the use of group decision support systems: The theory of adaptive structuration. In J. F. a. C. W. Steinfield (Ed.), *Organizations and communication technology*. Newbury Park, CA: Sage.

Poole, M. S., Hollingshead, A. B., McGrath, J. E., Moreland, R. L., & Rohrbaugh, J. (2004). Interdisciplinary Perspectives on Small Groups. *Small Group Research, 35*, pp. 3-16.

Raymond, E. S. (2001). *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary* Sebstapol, CA: O.Reilly.

Ridings, C. M., Gefen, D., & Arinze, B. (2002). Some antecedents and effects of trust in virtual communities. *The Journal of Strategic Information Systems, 11*(3-4), pp. 271-295.

Saunders, C., & Miranda, S. (1998). Information acquisition in group decision making. *Information & Management, 34*(2), pp. 55-74.

Scott, S. G., & Bruce, R. (1994). Determinants of innovative behavior: A path model of individual innovation in the workplace. *Academy of Management Journal, 37*(3), pp. 580-608.

Shaw, M. E. (1981). *Group Dynamics: The Psychology of Small Group Behaviour (3rd Edition)* New York: McGraw-Hill.

Shim, J. P., Merrill, W., James, F. C., Daniel, J. P., Ramesh, S., & Christer, C. (2002). Past, present, and future of decision support technology. *Decision Support Systems, 33*, pp. 111.

Shirani, A., Aiken, M., & Paolillo, J. G. P. (1998). Group decision support systems and incentive structures. *Information & Management, 33*(5), pp. 231-240.

Steiner, I. D. (1972). *Group process and productivity* New York: Academic Press.

Tung, L.-l., & Turban, E. (1998). A proposed research framework for distributed group support systems. *Decision Support Systems, 23*(2), pp. 175-188.

von Hippel, E., & von Krogh, G. (2001). Innovation by user communities: Learning from open-source software. *MIT Sloan Management Review, 42*(4), pp. 82-86.

von Hippel, E., & von Krogh, G. (2003). Open source software and the "private-collective" innovation model: Issues for organization science. *Organization Science, 14*(2), pp. 209-223.

Wikipedia. (2005, August). Wikipedia: The Free Encyclopedia. Retrieved August, 2005, from http://www.wikipedia.org

Wittenbaum, G. M., Hollingshead, A. B., Paulus, P. B., Hirokawa, R. Y., Ancona, D. G., Peterson, R. S., et al. (2004). The Functional Perspective as a Lens for Understanding Groups. *Small Group Research, 35*(1), pp. 17-43.

Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly, 25*(1), pp. 17-41.

Wong, Z., & Aiken, M. (2003). Automated facilitation of electronic meetings. *Information & Management, 41*(2), pp. 125-134.

Zhu, K., & Kraemer, K. L. (2005). Post-Adoption Variations in Usage and Value of E-Business by Organizations: Cross-Country Evidence from the Retail Industry. *Information Systems Research, 16*(1), pp. forthcoming.

Zigurs, I., & Kozar, K. (1994). An exploratory study of roles in computer-supported groups. *MIS Quarterly, 4*(1), pp. 277-297.