Email communications in team writing projects

Kirstie Edwards

A thesis submitted in partial fulfilment of the requirements of

Sheffield Hallam University

for the degree of Doctor of Philosophy

September 2006

Collaborating organization: Katholieke Universiteit Leuven

Acknowledgements

Trevor, Ben and Victoria made this work possible with their love, patience and support.

Ina, Keith and all my lovely family and friends have helped me, as always; thank you.

I also thank my supervisors. Simeon Yates coached me to overcome the hurdles in this last year, with a patient, listening ear. Geff Green gave me the benefit of his professional and moral support, with the advice of both a tutor and friend. Anne-Florence Dujardin taught me how to write for my readers, and she helped me to keep my feet on the ground, holding on to what I could realistically achieve. Noel Williams, my Professor, extraordinarily creative teacher and friend, opened up my mind to see the world in a different way. Thank you also to Professor Arthur Spaepen at the Katholieke Universiteit Leuven, for supporting me on my path towards a career in writing and research since 1991.

Thank you to the individuals at Namahn, Banksys, Keyware and the Society, who kindly consented to my researching their professional writing contexts. These contributors to my research showed openness, reflecting their professionalism, and for which I remain indebted. Many other colleagues from the Katholieke Universiteit have also supported me in my research, and although I thank them all, I am particularly grateful to Professor Martina Vandebroek.

I dedicate this work to Trevor, Ben and Victoria, who are my closest friends and family.

Abstract

Nystrand's social interactive model of writing (1989) describes composing as a social construct between the reader and writer. Although many researchers are debating whether the language of computer mediated communication (CMC) veers towards written or spoken discourse, or has developed a style of its own (e.g. Baron 1998, 2001; Ferrara et al. 1991; Harrison 2000a; Yates 1996), genre now plays an important role in fixity of documents in a digital medium (Yates and Sumner 1997). I therefore analysed the adaptations in linguistic 'style' of email writing as a non intrusive means to study social-interactive behaviour in networked team writing projects. Based on the premise that socio-emotional communications benefit team performance (Argyle 1994; Barker et al. 2000; Hyland 1998; Panteli 2004), I tested the concept that social interactive adaptations in email writing describe the social-task balance on projects and are reflected in the social interactivity applied to the team writing of the document.

Communication markers and writing influences were extracted from email content analysis to compare academic and commercial writing projects. Evaluating documents with Sless's social desirability model (2004) showed a parallelism between the social-task balance described by team emails and social desirability of the final documents. Further studies are required to prove this concept.

Social interactive adaptations demonstrated in socio-emotional behaviour in the emails of the academic project were also demonstrated in the final document. Higher pro-social behaviour was represented in Dutch emails in the academic project, and in English emails in the commercial project. Face to face contact influenced pro-social CMC behaviour and perceptions of behaviour.

The methodology provides a standard, unintrusive tool for monitoring the social dimensions of projects to identify and correct problem areas, and to research multiple contexts and inform more broadly on professional practice. Relating the social-task dimensions to document evaluations is the first step towards a causal model, to understand how team culture can influence virtual team writing.

The merging of personal and professional email styles predicted by Danet (2001a) is already apparent in communications from the academic context. Findings suggest that to encourage informal exchange of ideas and improve socio-emotional relations, professional email communications would benefit from a more conversational style.

Contents

Figures and Tables

Figures

Tables

1. Introduction

1.1 Thesis structure

This chapter outlines the thesis structure, research aims and research framework. In chapter 2, I review the literature on writing, virtual team and email research, introducing the rationales for my research and my epistemological stance. Chapter 3 describes ethical considerations and the methods used in the analyses of three professional team writing projects, which are reported in chapters 4, 5 and 6. Chapter 7 draws findings from commercial and academic case studies together to consider the influence of team culture, described by social and task dimensions of the projects, on social desirability (Sless 2004) of the documents. In chapter 8, I discuss the findings in the light of the literature and develop recommendations for future research and for professional email communication. Appendices A to G are included with the paper copy of this thesis and all appendices (A- Z and AA to LL) are included with an electronic copy of the thesis on the accompanying CD (see inside back cover). To avoid ambiguity, I have listed my intended meanings for some of the terms which are used throughout my thesis in a glossary at the end.

1.2 Research aims and rationale

My research attempts to answer the question: "Can we learn about the influence of team culture on virtual team writing from content analysis of email communications during projects?" By virtual or networked team writing, I refer to the professional collaborative writing practised by distributed teams in business, industry and academia. I describe team culture by profiling the balance between social and task dimensions on projects. The objective is to gain an understanding of the interactions between writing influences and social behaviour in networked team writing through analysis of email communications, to understand how team culture can influence virtual team writing.

The three hypotheses used to answer this question are:

> *Pilot study:*
>
> *H1 = Email communication behaviour is the product of writing influences and representative variables of both can be derived non-intrusively from email content.*
>
> *Main study:*
>
> *H2 = Social dimensions of teams can be identified from email communications.*
>
> *H3 = Social interactive adaptations and pro-social behaviour in a writing team's emails are reflected in the relational metadiscourse and social desirability of the document produced by the team.*

This research adds value to team writing research by combining different research perspectives, textual, individual, group and social, to study current, real virtual team writing practices, and by developing a standard research methodology for cross context comparisons. Results from studies completed using the same data collection and analysis techniques in different contexts

can be compared and used for meta-analyses, so that research findings may be more broadly applicable.

Additionally, the method includes the study of team dynamics during the process and outcomes in terms of participant feedback on the documents and document evaluations. The method thus contributes towards developing a causal model to understand how team culture can influence virtual team writing. Input variables are writing influences; these together with task and group maintenance activities during the process are profiled from email communications; the outcome variable evaluated is the social desirability of the documents. Data is searched for causal relationships between the writing influences, communication behaviours and social desirability of the document, to identify any hidden constructs or combinations of variables, which influence the team culture and the end product.

With application of the same communication behaviour to written email communications and to the end product of team writing projects, content analysis of emails provides a proxy method of research into professional writing practices. Analysis of email records retrospectively on writing projects provides an accurate, non-intrusive technique to study writing processes, without researcher intervention in the actual process or additional tasks introduced in the work context. Email analysis thus provides an ideal tool for the study of professional writing. My research therefore includes the design of such a tool to identify social constructs and group evolutions in writing research.

The originality of this research lies in the fact that it analyses project *email communications* for *team writing research*, an approach which has not previously been reported in the literature.

1.3 Practical framework and scope

The research comprises separate phases. In earlier research reported for my Master of Arts thesis (Edwards 2001), using email frequency on a writing project, I profiled the relative activities by task (purpose) and functional roles of individuals throughout the project. In the research reported here, I extracted indicators from the email content to represent writing influences, and communication behaviours representing characteristics of the group maintenance or social dimension of the project. Premised on the assumption that models of writing processes are applicable to written email communications in team writing projects, this initial pilot study addresses the following hypothesis, which is explained in more depth in section 2.4.

> *Pilot study:*
>
> *H1 = Email communication behaviour is the product of writing influences and representative variables of both can be derived non-intrusively from email content.*

I searched for dependencies between writing influences and socio-emotional communication behaviours, to understand how writing influences affect email communication behaviour in team writing in a commercial environment.

These dependencies were further explored in a second commercial project, for which perception data could be collected from the participants to construct meaningful interpretations. This part of the research aimed to show dependencies between writing influences and email communication behaviour in a second commercial project, and to develop the email analysis tool further to identify the social dimensions (see also section 2.4). A third case study used the same content analysis procedure to interpret dependencies in an academic community.

H2 = Social dimensions of teams can be identified from email communications.

Based on Nystrand's social interactive model of writing as a communicative act, we can draw parallels between adaptations required for context, audience and purpose of the final document and adaptations required for context, audience and purpose in communications to maintain the social structure of the group in team writing projects. The social dimension of team projects contributes positively towards performance, so the final hypothesis explores whether social interactivity and pro-social behaviour in team emails are reflected in the relational metadiscourse and social desirability of the document produced by the team.

H3 = Social interactive adaptations and pro-social behaviour in a writing team's emails are reflected in the relational metadiscourse and social desirability of the document produced by the team.

Social interactive adaptations and pro-social behaviour in commercial and academic projects were analysed together with evaluations of the final documents produced, focusing on social desirability of the documents, to explore how team culture can influence virtual team writing.

Figure 1-1 shows the overall research framework, to help explain the different phases, and also the type of data collection and analysis used in each part of the research. Different positions on knowledge, such as positivist, quantitative and scientific vs. interpretive and qualitative, influence how research questions are defined, choices of methodology and the knowledge acquired. To help validate my research, therefore, I adopt the stance of multiple realities of knowledge – that some is in the mind, some is outside of our minds and can be found, and some is created within and among minds – and adopt both qualitative and quantitative approaches. This research uses endogenic, interpretive and exogenic, positivist methods to support each other and avoid the bias of adopting a single stance. I discuss my argument for this in more depth in section 2.2.6.

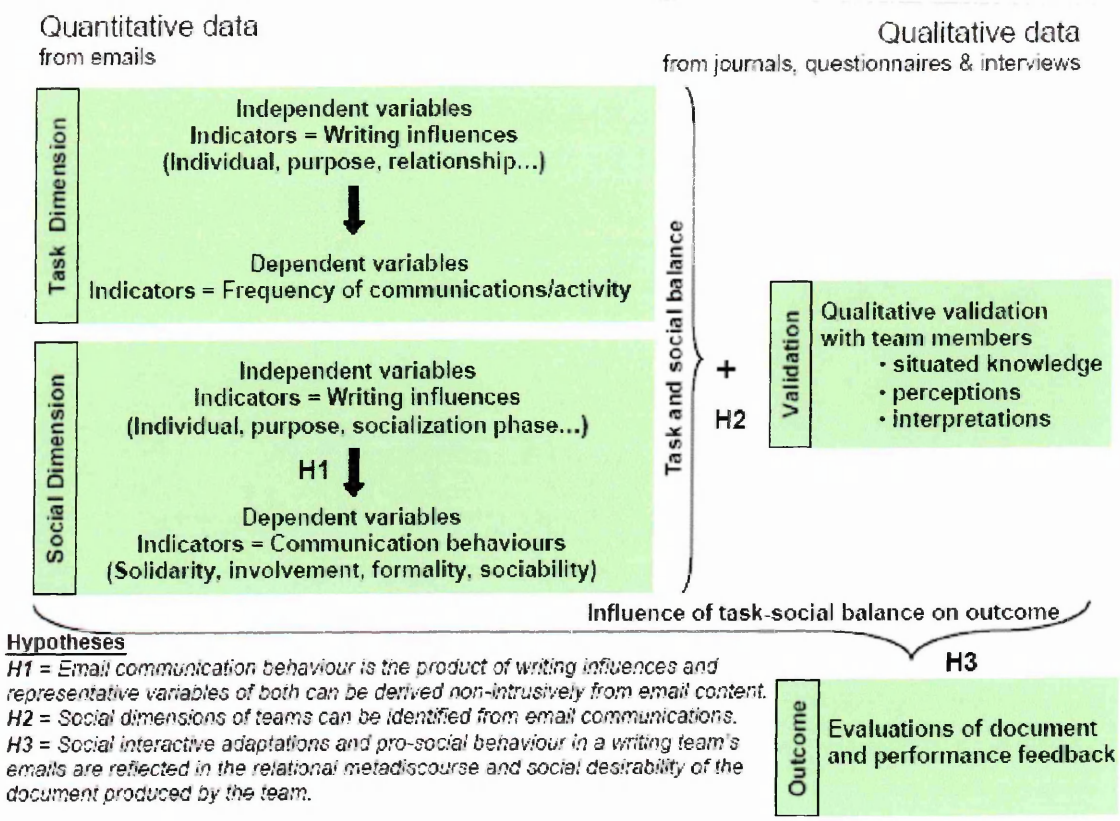

Figure 1-1: Research framework

2. Literature review

2.1 Chapter overview

This thesis reports research to answer the question "Can we learn about the influence of team culture on virtual team writing from content analysis of email communications during projects?" The research analyses *email communications* for *team writing research*.

Section 2.2 reviews writing research, analysing how writing research has evolved from the stages model to a collaborative writing model, and the choices of methodologies used. In section 2.2.6, I focus on epistemologies of writing research and my own epistemological stance. Relevant research on networked teams and computer mediated communication is reviewed in section 2.3. In section 2.4, I present the concept I explore in this research, the concept that content analysis of emails can describe social interactive team behaviour, which is reflected in the social desirability of the final document. Finally in section 2.5, I summarize the rationales for my research, which I have developed from the literature review.

2.2 Writing

2.2.1 Overview

My research analyses project email communications for team writing research. In the following sections, I review the changing trends in writing research relevant to research design, and briefly describe models which have contributed significantly to knowledge on writing processes and are most relevant to business writing: the stages, problem solving, social interactive and collaborative models of writing.

2.2.2 The stages model

This is a linear model with writing stages occurring sequentially: pre-writing (idea generation), writing (text generation) and rewriting (edit/reworking text). The focus is on the text as an external object. Writing is seen as knowledge transmission, a process of finding and structuring information and translating this information into words independent of the writer's ideas (Fitzgerald 1992 p18; Mitchell 1996 p9). Examples of early linear models include those of Rohman (1965 cited in Fitzgerald 1992 p18) and Britton, Burgess, Martin, McLeod and Rosen (1975 cited in Fitzgerald 1992 p18). In these models, writing was closely aligned with positivistic, exogenic viewpoints. Influences on composing were interpreted to be outside the writer, and the models targeted knowledge that writers needed about "universal textual attributes", i.e. external knowledge about what makes texts good (Fitzgerald 1992 p18).

2.2.3 The problem solving model

Outline of the problem-solving model

Flower and Hayes developed a model of the cognitive processes involved in a writing task from analysis of the verbal protocols of writers in action. This model brings a new element in that the sequential processes already identified in the stages model are intermixed. One activity may be embedded in another or be repeated at different times during composing: the model is recursive. "In particular, it specifies an organization that is goal directed and recursive, that allows for process interrupts and that can account for individual differences" (Hayes and Flower 1980 p29). The main components of the process were identified as:

- Planning: Sourcing information, organizing information and setting goals;
- Translation: Literally translating ideas into words, and
- Reviewing: Evaluating and editing the text, ideas and goals.

Internal to the process a production system, the monitor, orders sub-processes of planning, translating and reviewing. External influences on the system are the task environment and the writer's long term memory. The task environment includes everything outside the writer that influences the performance of the task, and therefore includes textual, individual and social factors: the writing assignment, intended audience, influences on the individual's motivation, and instantiated text. The writer's long term memory represents the writer's knowledge, not only on the content, but also knowledge of the audience and writing strategies.

The model from 1980 was further developed by Hayes in 1996, to differentiate in more depth between contributing characteristics of the individual, cognitive processes, working memory, long term memory and motivation/affect (Hayes 1996) and to show more clearly the interactions between the task environment (social and physical), and the individual. The model presumably includes the task implicitly in the task environment, but I would argue that the task, or the group goal should perhaps be a separate "input" entity (also discussed later in section 2.2.7).

This model may describe the cognitive processes in interpersonal email communication, where an individual writes an email. Although there is some debate over whether email communication style resembles conversation or writing (discussed later in section 2.3.6), Baron expects that as the traditional functions of letters are increasingly carried out by email in business and academic contexts, editing of emails will increase (Baron 2001 p242). We might reasonably assume therefore that the problem-solving model of writing applies to the writing processes involved in creating emails in professional contexts.

However, how important are these individual level cognitive processes to collaborative writing, if the resulting text is overruled by another individual after the process has finished, for example, if an idea that has been generated, translated and edited is eliminated by a co-author? A further research question is whether the model disseminates to group members, so that any process can be actioned by any individual in a team. To develop Flower and Hayes' model to

accommodate a collaborative context, where the sub processes are contributed by multiple collaborators, requires identification of newly introduced factors, such as the interface and interactions between collaborators, and knowledge of coordination and interpersonal skills etc.

Criticisms of the problem-solving model

The problem-solving model highlights procedural knowledge and skill and the ability to negotiate problems related to goal attainment.

> *Writing is viewed as a process of solving problems, that is writers have goals and purposes, and they act to achieve those goals. Problems arise when there are discrepancies between desired goal states and what is instantiated in text. The model is expressed as an informal information-processing model, which reflects endogenic/interpretivist tendencies. It focuses on the writer and allows for the belief that the mind can construct knowledge (Fitzgerald 1992 p20).*

Fitzgerald describes the model as a melding of exogenic and endogenic: "the constructive powers of the mind are given credence, but are seen from an outside observer's exogenic perspective" (Fitzgerald 1992 p29). She criticizes it for being centred on the writer, and cognition involved in procedural knowledge and skills. She also criticises how the model assumes that the knowledge of an author is static (exogenic), i.e. the knowledge on procedures for making decisions to solve problems is assumed to remain static. She reports critics as judging Flower and Hayes' "notions of mind" as reflecting logical positivism and ignoring the possibilities of "reflective, associative, metaphoric, intuitive, and imaginative thinking" (Petrosky 1983 cited in Fitzgerald 1992 p22).

Fitzgerald's strongest criticism, however, is in the choice of methodology which may have been influential in definition of the model. Using protocol analysis, which was already "typically used to identify processes in problem-solving tasks" (Hayes and Flower 1980 p3) provided a pre-conceived infrastructure for data collection, possibly influencing the results and interpretation. Indeed, Flower and Hayes write "we start with a basic premise: writing is problem solving, and can be analyzed from a psychological point of view of problem-solving processes" (Flower and Hayes 1981a p40).

Criticism of this probably most influential single work in the history of writing research is not lacking elsewhere. Nystrand criticizes Flower and Hayes' model because the least described part of the model is the translation of ideas to writing, and because it positions purpose as central (Nystrand 1989 p69). Much criticism of the problem-solving model is focused on its methodology, which while validated to identify psychological processes in other fields, had not been validated in writing research (Janssen et al. 1996 p233). Researchers challenging the use of think aloud protocols argue that thinking aloud may be reactive, i.e. "the writer's cognitive processes may be disrupted by the fact that they are writing and talking out loud at the same time" (Janssen et al. 1996 p233). Much think aloud analysis is based on the assumption that although it may slow the process down, it does not influence elements of the process or the outcomes, a theory proposed by Ericsson and Simon 1993 (cited in Janssen et al. 1996 p235). This theory was challenged by Russo et al. (1989 cited in Janssen et al. 1996 p236) as a result

of experiments which suggested that the possibility of interference could not be eliminated. Research by Janssen and colleagues testing this issue experimentally rejected two hypotheses that thinking aloud had no effect at all on writing behaviour and secondly, that it only lengthened the pause duration. From the rejection of these hypotheses, they concluded that the "reactivity of thinking aloud may very well vary with the writing task" (Janssen et al. 1996 p249). Other researchers question the assumption that subjects' short term memory content is in an orally compatible form for reporting. Cooper and Holzman (1983 cited in Janssen et al. 1996 p237) questioned the validity of the way in which protocols were elicited. Interestingly, a more recent methodology, Progression Analysis, is currently being used to analyse writing processes in the workplace (Perrin 2003, 2005). This multi-method approach combines interviews, participatory observations, keystroke-recordings, discourse analysis and cue-based retrospective verbal protocols. Progression Analysis appears to be a modernized protocol analysis for writing research, which removes the potential interference of writing and speak-aloud cognitive processes, the concern raised by Janssen and colleagues (1996) over the seminal work of Flower and Hayes (Flower and Hayes 1980, 1981a&b; Hayes and Flower 1980).

Hayes and Flower write that the "original model was derived through informal analysis of many protocols" (Hayes and Flower 1980 p27), and they provide example interpretations (Flower and Hayes 1981a). From the available sources (Flower and Hayes 1981a&b; Hayes and Flower 1980) it is not apparent how many students participated in the experiments, nor whether they were all from the same educational institution. Using subjects from one institution might introduce the influence of a common teaching methodology which becomes reflected in the model. Awareness of the research might also alter subjects' behaviour, for example knowledge that the research explores writing strategies might encourage a subject to strategize.

Although some researchers criticize the exogenic approach in trying to "corner" the workings of the mind in a positivist scientific manner, I would argue that the approach also has endogenic/interpretivist characteristics. Behaviours are interpreted and categorized. Here is an example: "Pat finally entered into a very productive session of brain storming, generating a series of ideas and examples that proved her point" (Flower and Hayes 1981a p54). From a social interactivism viewpoint, Flower and Hayes are part of the temporary shared social reality both in the verbal and written data collected, and thus are part of the system they research. They are the audience meeting the shared reality in understanding Pat's points and interpreting them as an indicator of a "very productive session of brain storming". This dilemma is the duality Fitzgerald (1992 p27) criticizes in the research. In claiming to take an exogenic/positivist viewpoint, (distanced from the subject and "cornering" thought processes) the researchers ignore their own involvement in the composing process, due to the missing facet of the social interaction and interpretivism. I discuss the point further in section 2.2.6.

Although the analysis was endogenic/interpretivist, Hayes and Flower (1980 p22-27) tested their model against three hypotheses in the analysis of a single protocol. Using verbal protocol data, they identified sections 1, 2 and 3, relating to the generating, organizing and translating

processes. Their model predicts that: generating is interrupted by editing; organizing is interrupted by generating and editing; and translating is interrupted by generating and editing. They tested whether items written in the different sections were logically characteristic of the processes identified from the verbal protocols: whether content statements in the protocol reflected the expected distribution of processes within the sections, and whether the generating process was most persistent in the section identified as such from the verbal protocol analysis. Distribution of characteristics of the writing in the different sections conformed to the model assumptions that generating would be most dominant in section 1, organizing in section 2 and translating in section 3, with an average inter-rater reliability of > 0.92 between 3 coders. Thus the form of written material varies corresponding to the changes in the process from section to section. These sections were identified from the protocol analysis; thus interpretation of the verbal protocols is reinforced by interpretations from the written text, which is evidence that the subject accurately reported what she was doing.

Their second hypothesis tested for evidence that generating appeared predominantly when the subject says the goal is to generate etc. and again the distribution of content statements was as expected, with 84.7% agreement between 2 coders. Testing segments taken out of context and coded by 2 further coders produced 67% and 77% coder agreements. Some of the discrepancy here may have been due to the difficulties in content interpretation of segments relating to editing, when taken out of context. Segments attributed to generating, organizing and translating had 86% coder agreement.

Finally by analyzing content ideas generated throughout the protocol, Flower and Hayes used number and linking of ideas (ideas cued by previous ideas), as an indicator of persistence in idea generation. Two coders had 96% agreement on linking, with 32 ideas in the generating section having an average linking of 6.4 and 16 ideas in sections 2 and 3 having average chain lengths of 2.

Thus, with reservations that this is only one protocol analysis providing empirical evidence confirming only three of the model's predictions, the results reinforce validity of the model. Flower and Hayes started with what they describe as "informal" analysis of many protocols and then in a mix of interpretive analysis resulting in numbers, at least partially justify their model quantitatively. Although they do not mention this in their article, these empirical studies also add credibility to protocol analysis as an effective means of evaluating what writers do, in the sense that they show that the verbal protocols match the written evidence.

Review of the problem solving model

The model developed by Flower and Hayes recognizes the recursive nature of writing, with planning, translating and revising activities interrupting each other, and also allows for individual differences. Flower and Hayes provide evidence to validate their model, which is widely accepted, despite the criticisms discussed above. This model may describe the cognitive processes in interpersonal email communication, where an individual writes an email. However,

to develop the model to accommodate a collaborative context where the sub processes are contributed by multiple collaborators requires identification of newly introduced factors, such as the interface and interactions between collaborators, and knowledge of coordination and interpersonal skills etc. The social-interactive model described next, provides the bridge to project recursive actions in this "problem-solving" and individual model, to a group context.

2.2.4 The social interactive model of writing

Outline of the social-interactive model

In the problem-solving model, composing is performed in a social context, whereas in the social interactive model of writing, composing happens as a social construct between the reader and writer.

Nystrand's social interactive model of writing (Nystrand 1982a&b, 1986, 1989, 1990) is based on the synthesis of research aimed at integrating social perspectives into our understanding of writing processes, including empirical studies of his own (Nystrand 1986). An important premise is the identification of "text as a communicative event" (Nystrand 1989 p73).

In this model the text is not autonomous, but a communicative event with a context of production and reception (Nystrand 1989 p73). The text is the negotiation of meaning between writer and reader. "We conceptualize text meaning, not in terms of the writer alone, but in terms of interaction between writer and reader purpose" (Nystrand 1989 p76). Nystrand describes the restraints writers experience trying to meet readers' purposes and expectations, as where social and cognitive factors interact in composing. He describes a "reciprocity based grammar" of written text, providing principles leading "to the flow of discourse between writer and reader" (Nystrand 1989 p80-81). These rules (which can be found in both Nystrand 1986 p80 and 1989 p81) outline the theory of reciprocity, the choices writers identify when reciprocity is threatened and the different types of elaboration the writer can use to redress the balance for convergence of reader and writer goals. Writers revise areas in their text where they feel reciprocity is threatened, i.e. where convergence with readers may fail, and the choices made involve either including or expanding on elaborations, perhaps with examples.

In this model, writing behaviour is influenced (among other things) by:

- *readers' expectations as the writer anticipates and interprets them;*
- *the impact of any previous communication with the reader;*
- *the effect of the text as the writer composes it on whatever remains to be written;*
- *any reader feedback that the writer anticipates;*
- *many characteristics of the context that give rise to the communication in the first place (Nystrand 1989 p75).*

Challenging the autonomy of text

Nystrand tested the autonomy of text empirically in a study in 1986 (Nystrand 1986 p81-107). Written and spoken communication vary in the sense that spoken communication happens in the same context as reception, whereas "written texts are composed for a context of eventual or potential use... determined by time,.... place...and purpose" (Nystrand 1986 p95). If written communication differs from spoken communication due to lack of writer's knowledge of the reader, necessitating autonomy of text, Nystrand argues that this should be demonstrable, through the tendency for written communication to be more explicit. Nystrand operationalized explicitness in terms of endophoric (referencing within the text) and exophoric (referencing externally to the text) references of written and spoken communications. However, in Nystrand's study, the frequency of endophoric references in the two communication modes did not vary significantly, indicating no more autonomy in texts than in the spoken word.

In the same experiment Nystrand also looked at different levels of abstraction; subject matter for communication was either reporting an event, or discussion of a speculative nature. Again there was no significant difference between the frequency of endophoric references between the two levels of abstraction. There were significantly fewer exophoric references, however, in the more abstract communication. He found that the ratio of endophoric to exophoric references varied significantly, from which he concluded that explicitness was a result of two effects, level of abstraction and mode of production (speech/written), which together show increases in endophoric and decreases in exophoric references.

Measurements of recall for the four types of communication showed no significant differences, although there were significant differences in the number of endophoric references among the four types. Nystrand argues:

> *This means that balance of understanding and overall coherence were generally maintained, but that several language samples achieved this coherence by means other than text cohesion and endophoric referencing (Nystrand 1986 p91).*

Nystrand believes that a "well-written text communicates not because it says everything all by itself, but rather because it strikes a careful balance between what needs to be said and what may be assumed" (Nystrand 1986 p96). Thus he argues that reciprocity underpins written communication just as much as spoken.

Reader-text interaction

A second experiment Nystrand reports (Nystrand 1986 p110-120) explores the Kintsch-Vipond reader-text interaction model, in which three criteria affect the interaction:

- Propositional density – a few ideas are easier to understand than many ideas in a fixed length of text;

- Number of new concepts per proposition – more information on one idea is easier to understand than a little information on many ideas;
- Text coherence – it is easier to understand related items.

Nystrand explored readability of texts revised to meet the three criteria above. One revision elaborated terms randomly, and another revised difficult lexical items, "those trouble spots where the writer's meaning and the readers' comprehension were not adequately matched" (Nystrand 1986 p113). Difficult words were identified in a process where words were left blank to be completed by respondents. Those words missed by 97% of the subjects were judged to be the "hard words". Revising randomly chosen words and the "hard words" according to Kintsch-Vipond's criteria (above) resulted in increases in recall of 12% and 60% above recall with the original text. Nystrand concludes that "text elaborations may not be random. Instead they must be carefully keyed to just those terms and concepts which are critical to readers' strategies of comprehension, to those terms that are more new than given" (Nystrand 1986 p118). Secondly he points out:

> *the extent to which writers' perspectives and readers' needs coincide or match up cannot be derived from even the closest analysis of text. Rather it must be based on a detailed assessment of the context of use – namely the needs, expectations and purposes of readers as they interface with what writers want to say (Nystrand 1986 p119).*

A recent experimental study on instructional text (Arts et al. 2004) has also shown that when writers are aware of the critical nature of the context of reading, over-specification increases. In this study subjects were requested to generate instructional language to identify buttons for different uses, one of which was surgical. "The results showed that language producers adapt their behaviour to the task to a high degree" (Arts et al. 2004), also confirming the situational influence of the reading context on the writing process.

Writer-reader interaction

An earlier empirical study of Nystrand's tested the writer-reader interaction hypothesis, that "the salient features of clear written communication lie... not in the interaction of the reader and text, but rather in the interaction between writer and reader by way of the text" (Nystrand 1986 p120). In this study, students were tested for IT knowledge and two groups, knowledgeable and unknowledgeable, were asked to read instructional texts and perform the related tasks. Subjects were permitted to ask as many questions as they liked during the tasks, and the questions were analyzed by two coders with 0.575 inter rater agreement. Differences between category and elaboration questions for the knowledgeable and unknowledgeable groups were significant, showing the trends illustrated in Figure 2-1.

a) Specification requests versus knowledge

b) Category requests versus knowledge

Figure 2-1: Proportion of questions requesting a) further specification and b) category definition v. knowledge of computer terminology (Nystrand 1986 p128-129)

The gist of Nystrand's interpretation of the results is that unknowledgeable subjects appear to be less concerned with detail and more with grasping the main theme, whereas the knowledgeable subjects have grasped the main theme and are more concerned with elaborating details.

A revision with higher topical elaboration significantly reduced errors made by the unknowledgeable group, but showed no significant reduction for the knowledgeable group. This supports the idea that topical elaboration improves unknowledgeable reader's comprehension – but does not affect the comprehension of knowledgeable readers, who already know the main idea – an observation Nystrand had made in an earlier pilot study (Nystrand 1986 p125).

> *Clarity of text obviously depends on the world knowledge that the reader brings to the text. Also writers compose clearly for their readers when they effectively take into account their readers' knowledge and expectations. ... assuring reciprocity between writer and reader may be said to be a fundamental motive for discourse production.... The extent and organization of the reader's knowledge is a critical factor ... and whether a particular text is ambiguous, abstruse or lucid depends on neither the writer, the reader, nor the text alone, but rather on the balance of all three - that is on how well the writer matches what he or she has to say, with what the reader expects and needs to know (Nystrand 1986 p132).*

The theme of "congruence between the discourse produced and interpretive frame of the receiver" had already been tested experimentally by Begg and Upfold (1980 cited in Nystrand 1986 p126). Clues were provided from one subject to another to help identify a target word. "Senders" were informed that the "receiver" knew or did not know the general category of the word (e.g. furniture). Whether this was true was manipulated randomly by the researchers. The results showed that the clues proved useful if intended for the correctly matched "receiver" (i.e. sender expected the receiver to have category knowledge and this was true, or the sender expected no category knowledge and this was true), but harmful in mismatch situations (Nystrand 1986 p126). Thus the accuracy of the "sender's" expectation of too much or too little category knowledge influenced the effectiveness of the communication.

Criticisms of the social interactive model

Nystrand's work is based on empirical research with statistical analyses and comprehensive analytical study of research elsewhere. However, his studies lack the "real-life" factor, being mostly experimental and using students as subjects.

Strauss (1987 p6-7) criticizes researchers for not addressing the complexity of reality. Writing is a social act (Faigley 1985 p236) and "social phenomena are complex phenomena" so the question is "how to capture the complexity of reality" (Strauss 1987 p10, 13). Many researchers (Carliner 2004a, 2006 p13; Diaper 1993 p92; MacNealy 1999 p4; Odell 1985 p250; Panteli 2004; van der Geest 1996a p9, 1996b p321-322) recognize the need for research into real professional writing practices. Panteli (2004) in particular calls for research into complete projects in professional environments. He writes:

> *...more studies are needed that collect and analyse the computer-mediated messages exchanged among team members who work on a joint virtual project. This will contribute to a better understanding of the dynamics of virtual teams and enable generalizability (Panteli 2004 p78).*

Thus we need to acknowledge the complexity of reality and explore data from real writing contexts, where all the influences are present and possibly changing, rather than artificially controlled.

Review of the social interactive model

Nystrand's social interactive model is premised on text being a communicative event and "hones in ... on the knowledge required for readers and writers to meet minds" (Fitzgerald 1992 p29) in the context of use. Nystrand believes that meaning of text is not autonomous, but that what writers write depends on expectations of the reader's knowledge, purpose and context of use (Nystrand 1986 p119). The model therefore tends towards an endogenic or even "interactivist" stance, discussed in section 2.2.6. Theoretically, the model provides a more comprehensive description of the writing process than the problem-solving model because it extends to include the reader context, relevant to both the writing of email communications and team writing. The model can be scaled from the interface of the individual writer and reader, to writing process *interfaces* between co-authors or writers and reviewers in a team. Indeed, collaborative writing may not only be perceived a good thing due to time saving and the combination of multiple skills (Edwards 2001 p57), but also because it brings usability testing, or the testing for reciprocity, within the writing process itself, rather than being an interaction that only happens after the document is finished. The iterative process in collaborative writing allows collaborators to review and comment on each other's contributions and negotiate to achieve the social reality, the reciprocity, the shared meaning. A model of collaborative writing is a social interactive model of multiple author writing, visualizing a shared social reality between writers and others, who may have multiple roles as readers, reviewers and writers. Developing the writing model further, Sharples describes a model of collaborative writing, which is discussed next in the following section.

2.2.5 Collaborative writing model

Outline of the collaborative writing model

The problem solving model depicts an individual cognitive process and the social-interactive process identifies composing as a social construct. From a more pragmatic viewpoint, Sharples models collaborative writing to include what writers actually do with the tools they have, based on much research towards a user-centred design of computer support (Sharples 1993 and 1996a). There are several main themes which Sharples newly introduces to our understanding of the writing process, including the role of external representations and the significance of communication and coordination in the process. These factors collectively embody a bridging of the cognitive and social processes in writing. In my research, emails are communications and external representations. Emails have permanence; they exist as records of idea translations to words and the sharing of ideas, the external representations necessary to achieve collaborative writing.

Sharples and colleagues found models such as the problem-solving model focussed on mental processes, not on the "operations, strategies and techniques carried out on some external medium" (Sharples et al. 1989 p26). To design a computer supported collaborative writing (CSCW) tool, Writer's Assistant, therefore, his approach was to explore user needs and writing processes in fundamental research (Williams 1992 p250). External representations are particularly relevant to collaborative writing because ideas may be shared and negotiated in the production process. Sharples refers to a broad spectrum of research on which the University of Sussex Collaborative Writing Research Group premised further observational studies of their own (Sharples 1996a p99). These latter studies recognize "the writer as a cognitive agent in a social and organizational context" (Sharples 1999a p7) and identify what writers actually do during writing tasks and what their needs are, with a view to accommodating their needs in the design of supporting technology (Plowman et al. 1993; Sharples and Pemberton 1988; Sharples et al. 1992; Sharples 1996a).

O'Malley's observational study of two subjects investigated the writing processes and use of external representations during true writing tasks (1988 cited in Sharples 1996a p100). The subjects were asked to "think aloud" and were video taped for later analysis while one worked on computer documentation and another on a thesis draft. Observations and analysis of the writing processes and how the subjects used existing resources, headers and meta-comments to shape their writing identified several important needs related to external representations. Subjects used existing resources as reminders, and annotations to plan and restructure their work and as reminders of work to be done, thus demonstrating how external representations coordinate the writing task.

In another study at Sussex, Wood filmed six pairs of students asked to generate and organize ideas for a short paper, using paper and board writing media, but no computers (Wood 1992). External representations produced fell into one of two categories, to store ideas and to support

conversation. Wood therefore interpreted the functions of these representations as to "mediate cognition" and "to mediate shared cognition providing the common grounding necessary to coordinate shared thought" (Wood 1992 p3; Sharples 1996a p102).

Beck's survey of 23 respondents identified the collaborative writing process as "dynamically renegotiated by the participants" requiring support not only for the work goal, but also the organization of the group and within group interrelationships (Beck 1993: Sharples 1996a p102). Two further case studies at Sussex also confirmed the need for support for communication, coordination, changing plans and external representations (1993 Plowman et al. cited in Sharples 1996a p103).

The needs identified in this research contribute to a perspective of the writing model not addressed in the problem-solving or social-interactive model. This new perspective includes how ideas become mediated, i.e. the actual practical bridge between the cognitive and the social, including representation of the translation of ideas, and their communication and organization. Sharples' categorizations of issues relating to collaborative writing originate from the objective of designing computer support for collaborative writing and are based on research into user tasks on and off line, and other models of writing. The categorization focuses on task, group, communication and external representation issues:

- Task issues include the work allocation and coordination, and interleaving of tasks. This he describes in strategies such as "sequential" where finite stages of the process move from one collaborator to another, "parallel" where collaborators can continue with separate work tasks simultaneously, and "reciprocal" working, where multiple contributors work on the same physical document simultaneously, perhaps accessing a common computer (Sharples 1993 p54). These stages in collaborative writing have been independently identified by separate researchers (Sharples 1996a p108), and may change throughout a writing project (Diaper 1993 p89).
- Group issues include role adoption, interdependence and management of conflict.
- Communication issues include context of communication, media influence and structuring or purpose of communication.
- External representation issues include types of representations used by writers, specifications of constraints, effects of media etc.

Sharples models writing as a cyclical process (Sharples 1999 p4) of planning, engagement, reviewing and reflection:

- Planning: planning what material to create and how to organize it;
- Engagement: creating the written material;
- Reviewing: re-reading the written material;
- Reflection: Forming and transforming ideas, "exploring conceptual spaces".

He describes this model as "not at odds with other problem solving models of writing (but).... stressing the cyclical recursive nature of writing" (Sharples 1999a p4). Writers may start with any of the activities and different writers may concentrate on different activities to the exclusion of others, thus accommodating both discovery and planner writers. Apart from these extremes on a continuum of writing approaches, which he describes as "Mozartian" (planner) and Beethovian" (discovery), Sharples also describes specific writing strategies identified, such as plan-draft-revise, outline-draft, draft-revise etc., (Sharples and Pemberton 1988 p5).

Criticisms of the collaborative writing model

In a similar way to each of the models already described, Sharples contributes a previously missing perspective to the picture, with the remainder of his model not conflicting with, but complementary to earlier models. Additionally the large body of research on which Sharples bases his interpretation of writing processes includes both academic and professional contexts and varied methodologies, such as case studies, surveys etc., which lends credence to his work. Beck surveyed professionals about their collaborative writing at work; O'Malley's observational studies were on real writing tasks, and Plowman's case studies included a student and professional writing context.

With the rapid and pervasive uptake of the Internet, people are increasingly expected to write collaboratively. Collaborative writing has also been encouraged by the growth of interdisciplinary studies. Even before the Internet became publicly accessible in 1992, surveys in non-academic writing in the late 70's and 80's confirmed the frequency of collaboration in writing work in professional settings (Anderson 1985a; Beck 1993; Lunsford and Ede 1990). However, recent survey research (Edwards 2001 p52) showed that 60% of the experiences reported related to difficulties encountered during collaborative writing, despite the fact that the practice of collaborative writing was perceived to be beneficial (60% respondents) in terms of time saved and the combination of skills. Negative experiences in collaborative writing are also reported in the literature, examples of which are competition between authors, anxiety over reviews and document ownership issues (Mitchell et al. 1995 p12-13; Petelin 2002 p178; Sharples 1993 and 1999b p169, 185). Thus, although there are strong indications that collaborative writing is a good thing, the process appears to have difficulties and there is scope for research aimed at improving practice.

Review of collaborative writing model

Sharples' model of the writing process broadly parallels the problem-solving model, but his list of influences on collaborative writing introduces the additional issues brought by group working, coordination and communication, with social issues contributing to all four of his categories, task, group, communication and external representation issues.

Rather than focus alone on the mental processes of single author writing, Sharples and his colleagues extended visualizing the process to include both pragmatics of the translating (i.e. turning concepts into text) and strategies for representing those concepts. Sharples' model

contributes the practical organization of the process. It takes a socio-cognitive approach combining the individual's cognitive processes with the organization of translation, representation, communication, negotiation and organization of ideas in a social context. Interestingly, email records provide a permanent representation of the sharing of ideas and coordination necessary in team writing, providing a form of external representation. These representations provide a means of studying writing processes in a consistent way across different projects for comparison of multiple case studies for the research of virtual and semi-virtual team writing projects. Such comparisons, together with project performance measures, will contribute towards identifying a causal model, to understand the influence of team culture on virtual writing teams. This concept is discussed further in section 2.4.

2.2.6 Writing research epistemologies

> *...some knowledge does reside in the text, some in the mind of the writer, some in the mind of the reader, and some is created through the interaction of all three —text, writer, and reader (Fitzgerald 1992 p139).*

Methodologies are founded on theories of language, learning and knowing (Fitzgerald 1992 p136); "to not see research methodologies as theoretical positions is to put the cart before the horse and to fail to understand how the instruments we use determine what we see" (Harste n.d. cited in Fitzgerald 1992 p135).

Research is a social practice in which individuals attempt to gain knowledge, either by creating or discovering it. Whether knowledge is created or discovered remains unanswered among other epistemological questions, such as "what is knowledge?", "where is it located?" and "how is it acquired?" There are abundant theories in response to these questions. However, Fitzgerald synthesizes these theories into three "world views" on knowledge, which logically underpin the reasons behind people's choices of methodology in research and interpretations of results. Two of these major views are exogenicism/positivism and endogenicism /interpretivism, and there is also a third philosophy emerging ("interactivist"), in which the methodology itself is instrumental and inseparable from the knowledge (Fitzgerald 1992 p4-7).

With exogenicism/positivism, also referred to as reflection-correspondence theory in cybernetic epistemology (Turchin 1991), knowledge is located in the real world and mirrored in the mind. Knowledge is found or discovered, and can be "cornered" using the right methods; "knowing" must occur objectively, with the knower distanced from what is sought. Knowledge is unchanging and exists independently of the knower. Knowledge is fact and a single truth. This matches the traditional scientific approach in research where only observable facts are interpreted as having credible existence and only objective research delivers credible results.

Endogenicism/interpretivism describes knowledge as originating in a person's mind, changeable, and unable to exist without the "knower". Knowledge may be facts, thoughts, feelings or emotions and there may be various versions of truth. Thus knowledge can be created and there are many ways of coming to know something (not an optimal way of "cornering" a single truth). "Knowing" may occur subjectively, with the knower working very

closely to what is sought. This mirrors the research methodologies used in social sciences, for example ethnographic studies in anthropology, to acquire knowledge of complex social phenomena, where it is impossible to control variables separately, or to simulate the real world scenario artificially for experimental purposes.

Fitzgerald writes that in the third and newly emerging view (von Foerster 1984 cited in Fitzgerald 1992 p9), knowledge emerges through interaction: "the observer, the observed and the process of observation itself form a totality, which cannot be decomposed into its elements" (Fitzgerald 1992 p9). The knowledge gained is therefore defined by the combination of the researcher, the knowledge itself and the means of acquiring the knowledge.

Writing research has evolved since the mid sixties from an exogenic viewpoint (knowledge exists and is discovered) where texts were analysed as stand-alone artefacts, towards an endogenic interpretation, where texts are recognized as a form of communication involving the originator, receiver and social contexts of production and delivery. Early research focussed on studies of the text, followed by the cognitive processes involved in creating the text, and finally the meaning of text as a temporarily shared social reality between the reader and writer. Naturally the social perspective has evolved with changes in social influences. With the increase in collaborative writing in business, industry and academia, (discussed in section 2.2.5), additional social factors such as group behaviour and coordination strategies necessary to accomplish team work have also been focused on in research. Finally, the increase in networked team working has attracted social research towards issues such as virtual team interaction and management in virtual collaborative infrastructures. The more traditional categorization of writing research (Faigley 1985 p233) from purely textual to individual and social, broadly speaking parallel Fitzgerald's epistemological categorization from exogenic, to endogenic and the third emerging "interactivist" view.

The textual perspective investigates the text as an autonomous entity, independent of its context. Measurable characteristics such as readability allow positivistic quantitative data collection techniques. Using the correct technique, several researchers can "discover" the same single "knowledge" answering the research question.

The problem solving model of writing, for which protocol analysis was used, (Flower and Hayes 1980, 1981a&b; Hayes and Flower 1980) depicts the mental processes involved in writing and appears to reflect exogenic and endogenic qualities (Fitzgerald 1992 p29). This type of research is interpretive and qualitative; the units of analysis are words. Although they attempt to describe procedures writers actually use, (Flower and Hayes 1981a p41) the researchers interpret the observed behaviours into categories to define their model. The model reflects endogenic/interpretivist tendencies, allowing for the belief that the mind can construct knowledge. However, writers' knowledge on procedures is interpreted as static, an exogenic quality. Fitzgerald criticizes the duality of their methodology: the researchers study the mechanisms of the mind taking partly an endogenic perspective, but also taking an exogenic

stance to "corner" the thought processes, while considering themselves distanced from the writer and the thoughts they want to examine (Fitzgerald 1992 p29).

Nystrand's social interactive model of writing also reflects the endogenic/interpretivist philosophy of knowledge. He claims that writers and readers interact every time the readers understand a written text (Nystrand 1989) and the goal is a shared social reality. His theory focuses on the interaction of minds, linking the reader, the writer and the text. Composing happens *as* a social construct in social interaction, whereas in the problem-solving model, composing happens *in* a social context. The model spotlights the linkages between readers and writers and the constructive powers of the mind, particularly as readers' and writers' minds interact. Fitzgerald identifies the "howness of mind meeting" (Fitzgerald 1992 p24), the knowledge creation and interchange between minds in Nystrand's theory as having parallels with the third evolving world view on knowledge, in which data collection methods are integral to the knowledge itself.

Through critical analysis of her own research, Fitzgerald (1992) shows how different positions on knowledge (for example positivist, quantitative and scientific vs. interpretive and qualitative), influence the research questions defined, choices of methodology and the knowledge acquired (Fitzgerald 1992). Different stances result in different choices and different knowledge being acquired. "When one method is used, complexities can be overlooked. When multiple methods are used and complexities emerge, the multiplicity can be used to re-enter, reconsider and re-examine the data" (Fitzgerald 1992 p139). The only approach to validate any type of research, therefore, is to accept multiple realities of knowledge, – that some is in the mind, some is outside of our minds and can be found, and that some is created within and among minds – and to adopt both qualitative and quantitative approaches (Fitzgerald 1992 p138). This research therefore uses endogenic/interpretive and exogenic/positivist methods to support each other, to challenge the bias of adopting a single stance.

2.2.7 Relevance of the models to this research

Overview

In this section I review the models already discussed and consider their compatibility and relevance to email communications and team writing.

The concern in modelling writing appears to be partly identifying the influences on the process and partly identifying the process itself. I first discuss the relevance of the writing process models to my research. I then consider the influences on writing from several different perspectives, which illustrate the complexity of researching real life practices, the need to study influences and outcomes to identify best practice guidelines, and the need for a standard methodology for writing research. Finally I discuss the changes in writing practices and their influences, which introduce the need for further writing research.

Writing processes

The models have areas of overlap, as might be expected as research progresses. Flower and Hayes (Flower and Hayes 1980, 1981a&b; Hayes and Flower 1980) focussed on cognitive processes happening in a social context. Fundamental to their findings was the issue that the process might not be linear, but recursive. Nystrand (1982, 1986, 1989, 1990) contributes significantly in extending the paradigm to the reader and the meeting of reader/writer minds through the meaning of text. Sharples and colleagues (Sharples 1993, 1996a&b, 1999a&b; Sharples et al. 1989, 1992; Sharples and Pemberton 1990, 1998) synthesize a cyclic version of the Hayes and Flower model, bridging the cognitive and social with practical representation, and introducing group, communication and coordination issues as influences in collaborative writing. The only discrepancy between these models lies in the focus of each. The problem-solving model focuses on goals and knowledge of writing strategies, whereas the social-interactive model focuses on a meeting of reader and writer minds. Sharples' model extends this to explain the movement between cognitive and social, the tools used, and the additional social needs in collaborative writing. These needs are the representation for communication and idea sharing, a function fulfilled by email communications.

Collaboration takes existing models of individual writing processes to a higher level, simply dividing tasks between individuals and introducing the additional interfaces and interactions required to communicate and coordinate with each other. Flower and Hayes identify the processes of planning, translating and reviewing in individual writing, which Baron expects to be increasingly applicable to email communications (Baron 2001 p242). In collaborative writing these planning, translating and reviewing tasks may be performed by different individuals. Collaborative writing introduces the reviewer as an individual other than the original author, simply bringing Nystrand's interaction with the reader (reviewer) within the writing process. The review is fed back to the author and the meaning negotiated in an iterative process until a meeting of minds or shared social reality is reached. The shared social reality is discovered through trial and refinement with writers and readers constantly changing roles and testing for what Nystrand calls reciprocity. Collaborative or team writing brings the writer-reader interface forwards in time, increasing the probability of a shared social reality, with the writers' text matching the reader's needs to achieve the common social reality of meaning. Group writing improves the outcome over single author writing, not only because it allows the collation of different knowledge and skills and may be accomplished in a shorter time (Edwards 2001 p57), but because it brings forward the reader-writer interface, allowing testing of reciprocity and iterative revision prior to usability testing of the end product, or real world use.

Writing requires not only the skill of wording intended meaning in an accurate and understandable way; it is also matching the representation of meaning to the audience's interpretation of meaning, and finding Nystrand's reciprocity. The writer tests the content for ambiguity, illogical argument, inappropriate order or presentation, and revises aiming to find reciprocity with the reader. The ability of an individual writer to achieve this reciprocity in an

interpersonal email, without the reader-writer interface being brought forward for testing of reciprocity (as possible in team writing), may have the potential to provide a useful means of predicting team writing performance. The coordination and communication necessary in team working, described by Sharples' model, and the representation of writers' communications in emails provide the means to research team writing using multiple case studies in a consistent way and allowing cross project comparisons. Based on Nystrand's social interactive theory of writing, content analysis of emails should identify adaptations in anticipation of readers' needs, adaptations also reflected in the final document produced in team writing projects.

The fundamental models already researched thus trace an evolution from narrower perspectives towards combined perspectives, both in terms of views on knowledge and different research perspectives, moving through textual, individual, social and group influences. The knowledge contributed by researchers towards an understanding of writing processes has built on previous findings adding to and reinforcing earlier interpretations rather than challenging them. I therefore consider that all of the theories discussed are relevant to my study of written email communications and team writing projects.

Writing influences

Professional team writing includes influences from all the perspectives, textual, individual, group and social, so writing research needs to combine these to model the process realistically. In an earlier literature review, I identified influences on writing and categorized them according to whether they were characteristic of the text itself, an individual, the group (including the collaborative process) or social factors (Edwards 2001 p12; see appendix A). Together with further variables relevant to group work, communication theory and interpersonal behaviour, I have re-categorized by group characteristics, individual characteristics, and social factors in appendices B, C and D respectively.

Individual characteristics centre on: knowledge, cognitive processing styles in writing, intelligence; human traits such as sociability, likeability, emotional stability and social competence; social roles functionally, organizationally, and in the writing context, and levels of socialisation within those roles; personal motivations, ideology, gender and age. (See appendix C for explanations and notes on these variables.)

Group variables reduce to 12 factors: purpose, group size, group age, group stability, membership composition (combination of individuals), role composition, cohesiveness, subgroup existence, norms, group status, group evolution and interdependencies. (See appendix B for explanations and notes on these variables.)

Social influences include all the influences, i.e. also those of the group and individuals. Additional influences are the writing context (or infrastructure of production), task specification, organizational and functional norms and their relative positioning, influence of other work on this work, previous texts and organizational goals. (See appendix D for explanations and notes on these variables.) Indeed, many researchers advocate the inclusion of social perspectives in

writing research (Adler 2000; Beaufort 2000; Faigley 1985; Mitchell 1996; Odell 1985; Te'eni 2001). Specifically in relation to email research, Duchenaut and colleagues write: "One simply cannot discount the organizational context in which the technology is used, the history of past interactions built over time, and anticipations of consequences for future interactions" (Duchenaut 2005 p33).

Studying influences in writing processes only has the potential to inform best practice if related to the outcome. Researchers such as Hart and Carliner are currently raising awareness of the need for standard productivity and performance metrics in writing (Carliner 2004b; Hart 2004). Research on the causal effect of combined individual, group and social influences on real-life writing practices is lacking. An understanding of the effect of combined variables on the process and results will identify predictors for optimizing group writing practices (Edwards et al. 2004, 2005). Assigning variables from the different research perspectives to stages in the input-process-outcome cycle is the first step towards data collection and analysis to identify a causal model.

From an ergonomic perspective any activity is an interaction between the task, a human and the environment. In group writing, the human is replaced by a group of humans, the environment by multiple environments networked together, and the task subdivided into multiple contributions by different group members. Each group member can interact with other members, and with sub-units of the whole task. The true environment becomes the mix of real and virtual worlds. Individuals interact with their own local environment and also with the overlying group environment, the locales of the team (Greenberg et al. 1999 p 32; Noel and Robert 2003 p246), which may be a mix of real and virtual, dependent on geographical locations of team members.

The process results from individual actions and reactions and is dependent on individuals' behaviour and anything that affects their behaviour. Situations, personality and the combination of situation and personality are influential on social behaviour (Argyle 1994 p102). The extent to which personality and situational (task and environment) factors determine social behaviour has been researched in a meta-analysis of 24 studies by Furnham and Japsers (1993). They found that the mean variance attributable to personality totalled 31.61%, to situation, 21.49%, and to the combination of situation and personality 46.9%. Despite the variations in the studies and dependent variables (e.g. cinema attendance or academic performance) in these different studies, Furnham and Jaspers concluded that the reanalysed studies provided strong support for the interactionist position (Furnham and Jaspers 1983 p638). In collaborative writing, individual characteristics (influencing individual behaviour), group characteristics (influencing group behaviour) and social factors all influence behaviour or interactions during the process. Thus the inputs to the process are the individual characteristics, group characteristics and social factors, and the latter include the environment (social contextual- and physical) and the task (or primary group goal). Table 2-1 categorizes input variables by group, task and environment.

Table 2-1: An ergonomic interpretation of input variables

Component	Influences
Group Characteristics	Group goal, group size, group age, stability, phases and evolution, membership composition (combination of individuals' characteristics such as knowledge and level of socialisation into a discourse community), balance, cohesion, subgroup existence, norms, group status and interdependencies.
Task Characteristics	Document specification, type (instructive, persuasive, commercial or academic) genre, purpose, content, intended audience, length, organization, deadline for completion, relation to other tasks.
'Environmental' Characteristics (Social and physical)	Media of cooperation (FtF, email, telephone, videoconferencing, audio/text connections), available technology, area of expertise, functional and organizational communities (e.g. academic, commercial, non-profit-making).

In section 2.3.2, I discuss how the activities in team work fall into two dimensions, which are social and task oriented. Some work aims towards achieving the goal and some activities maintain the group (coordinating, communicating, managing, maintaining good relations etc.). Variables influencing both these activities need to be included in any model of the process. Characteristics of these two activity types are listed in Table 2-2.

Table 2-2: Process activities

Component	Description
Task activities	Activities concerning the project content and striving towards the group goal. Actual task content such as the work of the group, the information it is sharing, the analysis it is performing the decisions it makes, the project it plans, interdependencies (Beck 1993; Sharples 1993)
Group maintenance activities	Activities concerning maintenance of the group and process, including management (work allocation, conflicts, ownership issues) (Beck 1993; Sharples 1993), communication (Faigley 1985) coordination and good relations (well-being of group members).

Finally, examples of outcome variables, i.e. process efficiency and effectiveness measures, are categorized in Table 2-3 (see also appendix E).

Table 2-3: Output variables

Component	Measures of success
Process Achievements	Success measured against individual, group and organizational goals Success measured against equivalent groups Perceived satisfaction of individuals, group, organization (Aytes measured perceived process satisfaction: Aytes et al. 2002)
Document Achievements	Usability of document: different methods are reported in Preston 2004 (this is equivalent to success measured against textual goals.) Reciprocity measures: Accurate audience interpretation of intended meaning (Nystrand 1989) / accurate writing of meaning needed by audience (response process model Hak 2004) Appropriateness for audience (Faigley 1985; Sless 2004) Appropriateness for audience context (Sless 2004) Topic progression and flow (Faigley 1985) Readability measures (Faigley 1985) (Flesch to be used with caution: Hartley 2004)

Although writing research has covered a diversity of methodologies, many of the findings cannot be generalized to a wider field, because they are based on artificial experimental- or single case-studies, or the results cannot be compared because the genres were too diverse (Carliner 2004a; van der Geest 1996a p9). We thus need to find techniques to collect and analyse data from different writing contexts in a standard way, which allows comparisons and broader applicability of the results. There is therefore still a need to develop a methodology which can be

used across different contexts in a standard way to allow valid comparisons of different studies in writing research. As such research may interfere with professional team goals in real writing contexts, the methodology needs to be non-intrusive. Such a methodology provides a platform from which to develop a model of collaborative writing, identifying performance predictors and informing on best practice principles. Content analysis of communications between co-authors in multiple case studies provides a solution to these problems for the research of virtual or semi-virtual team writing projects and is discussed later in section 2.4.

Changes in writing practices and influences

A final topic of relevance regarding networked writing teams and email use is the effect of advances in technology and communications on writing practices.

Much team writing research has been directed at perfecting the design of supporting technology or "groupware" (e.g. Cole and Nast-Cole 1992: Diaper 1993; Diaper and Sanger 1993; Dillon 1993; Gutwin and Greenberg 2002; Kim et al. 2001, 2002; Marca and Bock 1992; Sharples 1993, 1996a, 1999a; Sharples and Pemberton 1988, 1990; Sharples and van der Geest 1996; Sharples et al. 1989, 1992, 1993) rather than identifying how to improve team writing in whatever supporting environment exists. Diaper argued in 1993 that "what we need are models of the complex ways in which people wish to collaborate and these models should be derived from data where people are as unconstrained as possible by the technology they use" (Diaper 1993 p92). We do not need to focus on "what technology affords or permits us to do, but [rather on] ... how we appropriate the technology and make it do what we want it to do" (Thurlow et al. 2004 p51). This argument is now strengthened by the infrastructure offered by the Internet.

Internet uptake has been increasing since its first availability in 1992. By 2002, 90% of businesses with more than 10 employees in Europe were working on line (European Commission 2002). The Internet thus provides a backbone of supporting technology for collaboration across distance and time zones, extending the traditional means of collaboration to new ways of working together across networks, affecting the speed and nature of team work (Sharples and Van der Geest 1996 pv).

These changes brought by the advances in technology to writing environments and practices question the validity of applying traditional writing and team working theories to current day practices (Carliner 2004a). Carliner writes:

> *A combination of new technologies, new forms of communication, broader audiences and the changing manner of producing and using documents continually create opportunities to devise and test theories (Carliner 2006 p10).*

On the one hand, deficits of mediated communication are reflected in the media richness, lack of social context cues and social presence theories, and on the other, benefits are reflected in the social influence, deindividuation, social information processing and adaptive structuration theories (Thurlow et al. 2004 p48-51;p 66-67; see also section 2.3.4). Both benefits and deficits of mediated communication will have consequences on interactions and group dynamics in team writing practice, consequences which were not applicable in earlier writing research.

A further issue related to technology emerges from its distribution of the task environment in group writing and the social consequences (Edwards et al. 2005 p751). Faigley's concept of a discourse community being limited to members of an organization, or to an academic discipline (Faigley 1985 p238) is now more complex, and there are deficiencies and benefits to distributed and multi-disciplinary discourse communities, which are not experienced by co-located teams. Examples of this are the lack of opportunity for informal unplanned communications discussed in section 2.3.3 and the opportunities for absence afforded in virtual teams, discussed in section 2.3.4. Panteli writes:

> *...social boundaries in the traditional organization such as rules, procedures and control checks are used to control employees' productivity, whereas in the virtual organization different norms of conduct that emphasize output rather than input take precedence (Panteli 2004 p74).*

A short-term virtual discourse community may experience traditional social factors, together with additional influences, such as the removal of prejudices in the absence of visual and audio communication, the influence of anonymity (Barreto and Ellemers 2002) and isolation of remote workers (Larbi and Springfield 2004).

As supporting technology and working practices change, we need fast, accurate and non-intrusive methods of studying process dynamics and performance, which can encompass the working locales and contribute to knowledge of current day needs and practices. The new environment of team working, with new social boundaries, has new benefits and deficits, both defined by the medium and the distribution of groups, which I now move on to discuss in the next section.

2.3 Networked teams and email

2.3.1 Overview

The focus of this thesis is networked team writing, which happens in a mix of real and virtual worlds. In the previous section I have shown how writing research has evolved towards social constructivist theory, taking the social context, situational circumstances and relationships between writers and readers into account. In this section I review literature particularly relevant to my research, touching on team work in general, writing teams, virtual teams, and mediated and non-mediated communications.

2.3.2 Group dynamics and team performance

Overview

In this section I explain the relevance of longitudinal studies to researching behaviour in teams and define the two main activities in team work, which are task activities and socially oriented activities. Finally I draw on the literature to show the value of the social dimension in team working.

Development phases

While there appears to be general consensus that groups evolve through various stages, there have been several different paradigms proposed to describe these changes. Early work by Bales and Strodtbeck identified a three-phase model: orientation, evaluation and control (1951). In 1965, Tuckman synthesized previous research into the generally accepted 'forming, storming, norming and performing' model, and further updated this in 1977 to include a final phase, adjourning (cited in Gersick 1988 p10). Gersick critiques existing models for their lack of explanation of changes between phases, failure to identify how long different phases last and their treatment of the process as closed systems. She points out that "past research has concentrated on a few types of group and tasks, with little attention to naturally-occurring groups responsible for creating concrete products for outside use and evaluation" (Gersick 1988 p37). In a qualitative study of eight real group projects from beginning to end, Gersick was unable to match the dynamics to any of the existing models in the literature and developed a "punctuated equilibrium" model (Gersick 1988 p117).

> *A framework of behavioural patterns and assumptions through which a group approaches its project emerges in its first meeting, and the group stays with that framework through the first half of its life...At their calendar midpoints, groups experience transitions – paradigmatic shifts in their approaches to their work – enabling them to capitalize on the gradual learning they have done and make significant advances...Phase 2, a second period of inertial movement, takes its direction from plans crystallized during the transition (Gersick 1988 p32).*

In her discussion of this model, Gersick explains the early establishment of norms as being defined by material established before a group convenes, such as expectations about the task, expectations about each other, the context, individuals' behaviours and strategies. These factors influence the interaction in the first meeting. She describes the halfway point transition in real life working teams as a natural milestone, since teams have the same amount of time remaining as they have already used, which allows them to calibrate their progress (Gersick 1988 p34). The model also has two critical periods when groups are more open to influence; the initial meeting and the transition point. In the first meeting, "interaction sets lasting precedents; it holds special potential to influence a team's basic approach toward its project" (Gersick 1988 p35). Gersick also observed that the transition point was the only time when three conditions were met:

> *...members are experienced enough with the work to understand the meaning of contextual requirements and resources, have used up enough of their time that they feel they must get on with the task, and still have enough time left that they can make significant changes in the design of their products (Gersick 1988 p35).*

Gersick further reinforced her initial findings from the field study in 1988 with experimental research studying eight groups of students in 1989.

There are thus several paradigms for the phases of group development, but the constant factor is the dynamic nature; team behaviour changes over time. Researchers acknowledge the importance, therefore, of time-based studies, and this is highlighted below in section 2.3.5, in

the work of Walther and Chidambaram. Moving on from the chronology of team work, I next discuss the activities of the team.

Dichotomy of activities: task and social

Group communication and interpersonal behaviour theory together with writing research appear to clearly demarcate activities in team work into group maintenance tasks and goal oriented tasks.

Argyle describes two motivations for joining a group: "to carry out a task... and to enjoy social interaction and sustain relationships" (Argyle 1994 p167) and categorizes activities into task and maintenance dimensions. Argyle writes that in work-groups, "a pattern of informal social life develops as well" leading to group cohesiveness (Argyle 1994 p175). As early as 1948, Benne and Sheats classified leadership roles in team work into two types: task- and interpersonal-oriented. Bales (1951 p131) using his Interactive Process Analysis (IPA) noticed that groups typically switched between task and maintenance (social and emotional) activities. Slater (1965) in his study of 20 discussion groups differentiated between a "socio-emotional" leader and a task leader. Two profiles emerged from observations of activities and perception scores, a "best liked" person and an "ideas" person. "Ideas" people tend to specialize in active problem-solving attempts and "best-liked" people in more reactive less task-oriented behaviour (Slater 1965 p622). Further, in 1967, Fiedler developed a scale to differentiate between these leadership styles (cited in Hartley 1997 p98).

Having differentiated between the task and socio-emotional leaders in his study, Slater (1965) writes:

> *Since both are by definition highly valued in one way or another by the group, a high rate of interaction between them would be an indication that this relationship constitutes some sort of focal point in the group, and that the welfare of the group may be to some extent dependent upon the strength of this relation (Slater 1965 p620).*

Slater further explored the interactions observed to establish whether this might be the case and showed that the relationship was indeed "quantitatively important, although not always dominant in the group" (Slater 1965 p620). Study of the perception ratings of these two individuals compared to ratings of other team members also supported the concept that the relationship between these two was the most positive in the group. Slater concludes

> *We thus have the rather interesting picture of a respected task-oriented group member who is at best only moderately well-liked, receiving strong support from a perhaps more socially-oriented member who is the most popular man in the group, and with whom the task-oriented member forms a close and active relationship (Slater 1965 p621).*

Burgoon and Hale point out that "Despite the tradition of dichotomizing leaders and group discussants into task versus socio-emotional contributors, these do not have to be mutually exclusive categories. A person who is very task oriented may still demonstrate sociable tendencies" (Burgoon and Hale 1987 p40). At the level of interpersonal communication, an individual's skills in task and socio-emotional communication affect their interactions. At a team

level, from Slater's work, it appears that the combined contributions may affect the team's well-being.

Considering this social-task dichotomy in the framework of writing research, the collaborative writing model (see section 2.2.5) introduces the additional variables of communication and coordination. Coordination strategy in collaborative writing, division and organization of work, has been shown to affect workflow effectiveness (Dillon 1993). Beck describes from her survey of 23 collaborative authors, how discussions focused on content and structure of the document, organization of the work, and the relationships between the co-authors (Beck 1993). In her survey, the two purposes rated most important out of a list of seven options were firstly "getting the document done" and secondly, "to work together" (Beck 1993 p101).

Cole and Nast-Cole (1992) differentiate between the two activities in team work as follows:

> *Task activities are often what is thought of as 'work' in organizational settings and are directly related to the job at hand or the purpose for the group's existence. Maintenance activities, on the other hand, are often 'invisible' in work settings and are those activities that focus on the well-being and solidarity of the group (Nast and Nast-Cole 1992 p48).*

There appears to be general consensus across disciplines, therefore, that team work requires both task and socially oriented activities, and I now discuss the influence of the social dimension on performance.

Social dimensions and team performance

Team performance increases to an optimum with group cohesion and then decreases (Argyle 1994 p168; Evans and Dion 1991; Kelly and Duran 1985; Root 1988; Wilson 1986 p243), suggesting that factors increasing group cohesion, such as pro-social behaviour, may be as relevant to performance as professional skills (Argyle 1994 p156; Barker et al. 2000; Hyland 1998 p241; Panteli 2004 p76). I discuss here evidence from the literature suggesting that the balance between task and socio-emotional dimensions in teams affects the welfare of the group.

Describing professional contexts as social institutions where tasks are carried out, Root (1988) recognizes the critical elements for success are interpersonal communications and informal social relationships. In contrast to traditional approaches to researching design of computer supported cooperative work, Root (1988) therefore focused in his research on the design of tools to support unplanned, informal social interaction, a valuable aspect of co-located groups which distributed groups miss (see the discussion of Kraut's work in section 2.3.3).

Kelly and Duran (1985) studied the cohesiveness and performance of seven groups of students working on problem-solving tasks for which they had to prepare an oral and written report. They defined group cohesion as "the extent to which members of a group stick together, like and respect one another and feel unified" (Kelly and Duran 1985 p186), thus reflecting the social dimension of team work. Using Bales and Cohen's adjective rating method, they derived a group average score from member scores of perceived behaviour rated on three dimensions:

dominant vs. submissive, friendly vs. unfriendly and instrumental vs. emotionally expressive. They interpreted group cohesion from close clustering of scores for members' perceptions of one another and themselves on these dimensions. The researchers categorized the average instructor gradings for the oral and written reports provided by the groups and reduced these to high and low categories. Based on calculations of average Euclidean Distance scores for interpersonal distance (average distance among group members on the three adjective rating dimensions), they found that groups with either very high or low scores, equating to low or high scores for group cohesion respectively, did not perform well. The researchers therefore tentatively concluded that an optimal level of group cohesion might exist, cautioning however, that the study was small in scale.

In 1991, Evans and Dion completed a meta analysis of 16 research studies examining group cohesion and performance. These studies focused on team work for which performance is easily measured, such as sports, so that the results may not be applicable to real work groups such as virtual writing teams, in which performance is less overtly measurable. The researchers also point out that results may be influenced by methods used to assess group cohesiveness and by the retrieval bias of only studying published research. However, their results clearly suggest a positive relationship between group cohesion and performance.

Mortensen and Hinds found in their study of 24 teams that shared identity was significantly associated with performance (Mortensen and Hinds 2001 p231), and communicative predictors have been shown to affect virtual team outcomes in terms of perceived cohesiveness, trust and satisfaction (Timmerman and Scott 2006). Tucker and Panteli (2003) studied global virtual teams in a high tech multinational organization and found that the teams which worked well included a "social and fun element in their computer-mediated interactions, which appeared to have helped in creating stronger relationships (Panteli 2004 p76; Tucker and Panteli 2003 p95).

There is therefore strong evidence in the literature, not only that a social dimension exists in team work, but that it contributes in a positive way to the performance of the team.

Review of group dynamics and team performance

In this section I have discussed three aspects of team processes and these are the dynamic nature of team behaviour, the dichotomy of task and social activities, and the positive contribution of the social dimension to team performance. This last issue is particularly relevant to distributed writing teams communicating by email, because some theories of mediated communication suggest that email exchanges may not support socio-emotional communication as effectively as FtF exchanges, and I discuss this further in section 2.3.4. First, I step back in time a little to review a benefit of co-located teams.

2.3.3 Non-mediated communication

Overview

In the previous section, I highlighted the dynamic nature of team behaviour, two types of team activities, task and social, and the value of the social dimension to team work. In this section, I review some early work by Kraut, Egido and Galegher, (1988) emphasizing the value of informal communications in co-located teams, and the lack of such opportunities in distributed groups.

Physical presence and opportunities

Kraut and colleagues completed a very interesting study in 1988, on the collaborations between scientific researchers. Their research identified the value of informal interpersonal communication in building collaborations, and informed on the social interactions missing from dispersed collaborations, rather than on the restrictions of a particular medium used for interactions. This work didn't focus on the leanness of email or telephone communications, but rather on the lack of spontaneous meetings afforded by collocation.

Studying 4278 unique co-author pairings among 93 researchers, who had published at least two internal research reports in 1986 and 1987, they collected data on organizational proximity, physical proximity and research similarity. Using logit analysis, they held organizational proximity and research similarity constant, and showed that physical proximity has an independent effect on research collaborations. The likelihood of collaboration was higher between scientific researchers who were physically located close to one another than between those on different floors or in different buildings, and the researchers attribute this to the frequency of communications. "The informal contact that results from frequent opportunities for communication often leads to collaboration" (Kraut et al. 1988 p5).

Further, in a survey of collaboration among psychologists, Kraut et al. asked the respondents to indicate the distance between their offices and those of the primary co-author for each of their collaborative articles, and to estimate the frequency of their communication with this co-author. The results demonstrated that physical proximity is strongly related to frequency of communication during both the planning stage and writing stage of the research process. There may be a slight bias here in that physical proximity may influence the perceptions of respondents in estimating frequency of communications, which Kraut et al. do not appear to have taken into account in their discussions. However, the results suggest that co-authors with adjacent offices communicated twice as often as those pairs who were simply co-located on the same floor. Co-location increases frequency of interaction, which in turn increases the likelihood of collaborators liking each other and therefore of further collaboration (Kraut et al. 1988 p6).

When Kraut and colleagues published this research paper, they were able to draw on earlier research showing that the phenomenon was not restricted to face to face (FtF) communication. They present the results of earlier research by Allen in 1977, whose study focused on industrial research and development engineers, and showed a logarithmic decline in FtF communication

frequency with distance between potential communicators. Mayer (1976 cited in Kraut et al. 1988) had shown decreasing communications mediated by telephone with increasing distance between collaborators, and Eveland and Bikson (1987) had shown the same for communications mediated by email.

Surprisingly in the study by Eveland and Bikson (1987 p97) of over 69,000 emails at the Rand Corporation during the first stages of the email system's implementation, 45% of the emails were sent to other people in the immediate physical location, which is an interesting result contradicting the concept that email is primarily for communication with distant people. This together with the negative association between email frequency and distance led Eveland and Bikson to conclude that email improves existing interaction, rather than initiating new interaction patterns. The 18 month study showed no evidence of the formation of non co-located collaborations (Eveland and Bikson 1987 p100). Additionally for the most frequent users, email appeared to become the normal means of communicating rather than a way of overcoming distance.

The fall in communication frequency with distance led Kraut and his colleagues to conclude that "much communication between actual and potential research partners is not planned and would not occur if it had to be planned" (Kraut et al. 1988 p9). They further analysed interview data to understand the effect of quality of communication on collaboration and concluded that the "informal communication is important because it allows researchers to develop common interests with their neighbours" (1988 p7). "Current communication technology available to most researchers does not allow the intensity of interaction or the spontaneous exchange of notes and documents that are typical of the FtF meetings" (1988 p7). Further, Kraut and colleagues argue that proximity "provides a low-cost opportunity for a researcher to discover the qualities of another that might make him or her a desirable collaborator" (1988 p8). Most of the researchers they interviewed supervised subordinates and coordinated with peers during casual hallway and lunchroom conversations, just as often as during formal scheduled meetings. The researchers therefore emphasize not only the cost restrictions on travel and telephones to maintain dispersed collaborations, but also how "'on the fly' interactions are impossible in collaborations that occur over a distance" (1988 p8). They conclude several opportunities afforded by proximity, including people being able to inexpensively and informally assess how well they might work together. Once committed to the collaboration, frequent low cost communications allow collaborators to chase work and report to each other informally through casual interactions, and to share decisions and develop a sense of co-ownership of the work.

> *Finally, and perhaps most importantly, throughout the collaborative process as a whole, proximity supports a convivial personal and working relationship by building a consensus of views and interests and maintaining shared knowledge about the project and about the local culture in which it is embedded (Kraut et al. 1988 p9).*

These researchers criticize tools designed for collaborative working because they only target formal planned communications. From their recognition of the value of interpersonal communication other than those which are specifically scheduled and task-related, Kraut et al.

identify three tools for team work and especially for distributed teams: 1) communication tools to facilitate both planned and unplanned real-time and delayed interactions 2) coordination and management tools and 3) task-oriented tools.

Now, with cost restrictions to virtual working removed, the question remains whether social presence in virtual working can match the richness of that in the real world, not for example because social cues are missing from lean email media, but because the communications are wilful and task oriented, rather than coincidental and unplanned. "Many of the interactions that make up this feedback over time are damaged by intentionality and simply would not occur if they must be wilfully initiated" (Kraut et al. 1988 p9). An interpersonal communication will not arise by chance in the virtual world as a result of other human activities, as it might in the real world.

One final aside to Kraut and colleagues' observation that communication frequency fell with distance is that since 1988 costs of CMC have become less restrictive to professional practice in developed countries. This economic evolution, however, does not detract from Kraut's observations or deductions. In 2001, Mortensen and Hinds reported a survey of 24 teams from 5 companies, 50% co-located, 29% domestically distributed and 21% internationally distributed. Interestingly they found no indication that distributed teams used communication technologies more than co-located teams. Co-located teams reported that over 50% of the communication was mediated. Thus still in 2001, CMC did not appear to increase with distance, as we might expect. The informal and unplanned communication opportunities afforded with collocation, which help develop relationships, may also have a knock-on effect on the likelihood of communicating by CMC. Further, Mortensen and Hinds' findings suggest that using email content to study team writing projects may also be applicable to co-located teams, if sufficient email records are available. Most important to my current work, however, is that research aimed at designing supporting technology for collaborative writing had already identified the need to support informal interpersonal communication as early as 1988.

Review of the value of non-mediated communication

Before I move on, I will highlight again that the main theme emerging from the work of Kraut et al., which is relevant to my research, is that proximity provides opportunities for informal unplanned communication, which have no equivalent in email communications. An email I send to a colleague in India, will offer me no new opportunities for interpersonal contact with the person sitting next to that colleague. If, on the other hand, I walk next door to speak to my colleague, an opportunity may well arise for me to speak to the person sitting next to that colleague. The research by Kraut and colleagues, which I have described here, suggests that this type of unplanned informal communication is extremely valuable to team work. Textual, computer mediated interactions such as email, however, preclude such opportunities. The lack of opportunities for informal interpersonal communication remains a separate issue to the 'deficit theories' (Thurlow et al 2004 p48) of CMC discussed below.

2.3.4 Mediated communication

Overview

> *...in so far as language is not a system of communication on its own, but part of a larger semiotic web of intended and unintended signs and phenomena, the sharing of verbal communication alone may curtail or filter the sense we have of others, by altering what about people and the way they live is available to us for interpretation (Rooksby 2002 p5)*

In the previous section I discussed the opportunities for communication afforded by physical presence. In this section I discuss theories which try to explain the benefits and deficits of mediated communication and which relate to what the communication quality, rather than opportunity, allows us to learn about the writer, writer's context, and writer-reader context. I discuss seven theories: media richness, lack of social context cues, social presence, social influence, social identity and deindividuation, social information processing theory, and finally adaptive structuration. Although each of these models takes a slightly different approach to explaining mediated communication, all consider how sharing text alone affects the information available as Rooksby describes above, either by reducing information, or by leading to interactive adaptation and development of alternative communication strategies.

Theories and research

Media richness theory (Daft and Lengel 1984) focuses on the bandwidth or number of cue systems available within different media. FtF is considered the richest medium due to the availability of immediate feedback and the number of channels used. CMC is leaner because no nonverbal cues are present. Media richness theory suggests that communications of potentially ambiguous information are most effective with rich media and less ambiguous information can be communicated by leaner media.

An early content analysis of 157 emails received by a middle level manager (Sherblom 1988 p49) showed a higher frequency of mails designed to exchange information than for the more complex communicative functions, such as personal, social and influence attempts. However, later experimental research (e.g. Dennis and Kinney 1998) has shown that matching media richness to task equivocality does not improve decision-making performance. Lynn McGee (2000) has analysed media choices in real professional writing environments and found that decisions could not be explained by media richness theory alone.

Sproull and Kiesler's (1986) lack of social context cues theory differentiates between different media, FtF or CMC, on the basis of the social information available.

> *Once people perceive social context cues, these cues can create or elicit cognitive interpretations and concomitant emotional states. People adjust the target, the tone and verbal content of their communications in response to their definition and interpretation of the situation. Typically, when social context cues are strong, behaviour tends to be relatively other-focussed, differentiated and controlled. When social context cues are weak, people's feelings of anonymity tend to produce relative self-centred and unregulated behaviour. That is, people become relatively unconcerned with making a good appearance...Their*

behaviour becomes more extreme, more impulsive, and less socially differentiated (Sproull and Kiesler 1986 p1495-6).

In their study of 1248 organizational communications from a Fortune 500 company, Sproull and Kiesler (1986) explored the effects of the medium on self-absorption, status equalization and uninhibited behaviour. They studied the contents of actual emails, and self-reports for comparison of email behaviour with behaviour of other media. Their findings confirmed the relatively weak social context cues in emails and that decreasing social context cues has substantial deregulating effects on communication. For example, people focused more on themselves than on others in email greetings (self-absorption), emails written by superiors and managers did not vary from those written by subordinates and non-managers (status equalization), and people behaved irresponsibly in emails more often than in FtF interactions (uninhibited behaviour). (I return to this deregulating or decreased inhibition concept further in my discussion of research on socio-emotional exchanges in CMC in section 2.3.5.)

Social presence theory evolved to describe teleconferencing and is frequently applied to CMC phenomena. Walther (1995) defines social presence as:

> *the feeling one has that other persons are involved in a communication exchange. The degree of social presence in an interaction is posited to be determined by the communication medium: the fewer the channels or codes available within a medium, the less attention is paid by the user to the presence of other social participants (Walther 1995 p188).*

Thurlow and co-authors (2004) define social presence as "the level of interpersonal contact and feelings of intimacy experienced in communication" (Thurlow et al. 2004 p48).The social presence theory suggests that having fewer visual cues, (e.g. facial expression, posture, dress etc.), leads to low social presence, which in turn leads to more task-focused and less relationship-focused communication. Email communication would therefore be expected to invoke low feelings of social presence, in comparison to FtF meetings. Sherblom (1988), in his earlier research into email communication, anticipated a change in organizational communication as a result of using email. He expected more emphasis on information exchange and less on personal social and negotiated communication, and deficit theories such as the social presence theory describe this concept.

However, social influence theory suggests social rather than technologically driven determinants of media choices (Schmitz and Fulk 1991). Schmitz and Fulk premised their research on the belief that "social interaction in the workplace shapes the creation of shared meanings and that these shared definitions provide an important basis for shared patterns of media selection" (Schmitz and Fulk 1991 p488).

Social influence theory proposes that:

- media perceptions and use are partly socially constructed;
- media properties are subjective (and therefore individual), partly influenced by attitudes, statements and behaviours of others in the workplace and also by individual differences in medium expertise;

- media choices are not necessarily efficiency motivated;
- media choices may be designed to preserve or create ambiguity to achieve strategic goals (Schmitz and Fulk 1991 p491).

In their study of perceived media richness and social influences from organizational colleagues on the uses and assessments of email in a large research and development organization, Schmitz and Fulk analysed 511 questionnaires from email users and completed post survey interviews with 27 respondents. They tested a number of hypotheses which predicted that:

- experience in use of email, computers and keyboards would predict perceptions of email medium richness, which in turn would predict reported email use and perceived usefulness;
- supervisor and co-worker perceptions of usefulness and reported use would predict an individual's reported use and perceived usefulness of email.

Both keyboard skills and computer experience influenced richness perceptions, showing that medium expertise affected perceived medium richness. The hypotheses linking close communication partners' assessments and use of media to an individual's media assessments and use were also generally supported (Schmitz and Fulk 1991 p513).

McGee (2000) completed a survey of 30 technical communicators on the communication channels chosen during documentation projects. She analysed the data against media richness and social influence theories to understand the reasons behind the choices made. FtF communication was preferred at the beginning of projects, and email towards the end, and for managing conflicts. The lack of interpersonal elements in lean media was an advantage when respondents needed to focus on the facts and when there were interpersonal problems. As an example to illustrate how media richness and social influence theories might be tested to explain choices, McGee describes a disagreement between two managers at Microsoft who chose to use email to avoid heated FtF interaction, and to resolve a conflict of opinion (McGee 2000 p38). In this example, media richness theory would predict that email is not the optimum medium for the ambiguous task of conflict resolution. In reality the leaner medium was chosen to "diffuse the emotional content of their previous interactions, while still salvaging their ability to continue working on the same project" (McGee 2000 p38). Applying social influence theory, McGee suggests that email was promoted and routinely used at Microsoft, and that another organizational norm, the work ethic to resolve conflicts which hinder project progress, might have played a role in the medium choice. The example may also illustrate two propositions of social influence theory, that media choices are not necessarily efficiency motivated and may be designed to preserve ambiguity, for example in this case, to preserve ambiguity over an emotion.

In concluding on her survey data, McGee reports that media richness and social influence theories helped to explain many of the respondents' media choices, suggesting that these theories may be complementary. However, she also discusses the possibility that media

choices in professional practice may fall outside either the media richness or social influence theoretical frameworks, writing "comments from these respondents indicated very consistently that they choose 'whatever works'. By that I mean that they choose the communication channel that will most effectively get the job done at the time" (McGee 2000 p48). Thus the particular task and context influence an individual's choice of medium, rather than predominantly the appropriateness of the medium for the message type (media richness theory) or what is expected as normal behaviour in a particular organization (social influence theory). McGee's results, therefore, require a more social interactive theory, and may support adaptive structuration theory (discussed further below), in which interactions, rather than technology or individual attributes, are influential.

The 'Social Identity model of Deindividuation Effects (SIDE)' proposes that factors traditionally identified as causing deindividuation, such as anonymity and group immersion, or interaction with a computer network, can actually reinforce group salience and conformity to group norms. With the tendency to move from individual to social identity in virtual contexts, and to give people the benefit of the doubt in the absence of social and nonverbal cues, social identity becomes stronger and contributes to greater cohesion (Postmes et al. 1998; Thurlow et al. 2004 p67). Postmes and colleagues explain that the cognitive aspect of the SIDE theory proposes that where it is more difficult to represent the individual, as in lean communication media, sensitivity to social norms is increased (Postmes et al. 1998 p698). These authors continue to explain the paradoxical implications of this in mediated communication:

> *CMC is not necessarily the impersonal and businesslike medium it is so often portrayed to be. Rather the medium can be perceived as a socially rich environment, in which available cues to a shared social identity gain great weight due to the absence of individuating information (Postmes et al. 1998 p698).*

This theory is rather analogous to visually impaired individuals developing and optimizing unimpaired senses.

In Burke and Chidambaram's research (1999) comparing FtF and mediated communications in an experimental setting, performance in synchronous groups was significantly better than FtF groups. With the deindividuation theory, we might expect higher social identity in the synchronous groups, which by providing the social dimension to the team work contributes positively to team performance. However, Burke and Chidambaram explain the improved performance in the synchronous group as a result of the lack of relational communication, and the ability to hypercommunicate with lean media, i.e. to focus precisely on the task (or message) and to control and select specifically how and what to present (Burke and Chidambaram 1999 p572). Thus whereas lean media are predicted by media richness theory to be less supportive to ambiguous tasks, these researchers argue that the lack of relational or socio-emotional communication makes them *more* supportive of such tasks. Thus Burke and Chidambaram ignore the added value of the social dimension identified in team work, rather interpreting a lack of socio-emotional exchanges as beneficial. (The nature of mediated textual communications in this research providing permanent records may also have contributed to the success of the

teams using mediated communications, helping them to remain task-focused and aiding representation, sharing, retrieval and revision of subject ideas. However, a third asynchronous group did not perform significantly better than the FtF group.)

Walther's social information processing theory proposes that *rates* of socio-emotional exchange differ between FtF and computer supported groups; given time, people learn new ways of "verbalizing relational content" (Thurlow et al. 2004 p51; Chidambaram 1996 p143). Walther acknowledges that people communicate in professional organizations to achieve tasks, but also that as "they do so, they also communicate to manage their interpersonal identities, their roles, as well as the character of their relationships with others in a process known as relational communication" (Walther 1995 p186). He emphasizes the importance of relational communication to job satisfaction, but argues that the deficit theories, social presence, cuelessness and media richness theories, assert "that the structure of the medium alters the nature and interpretation of messages, it implies that such effects are inherent and constant whenever people communicate using computers" (Walther 1995 p188). He points out that this precludes the influences of relationships, context or dynamics, such as development and changes occurring with time. This theory has much in common with the social influence theory, taking a more interactive stance.

Early confirmatory research supporting the deficit theories has often been reversed in field studies or longitudinal research (Walther 1995). To address contradictory theoretical specifications and empirical findings, Walther analysed relational communication in 32 three-member student groups, half of which were CMC and half FtF. Groups had to complete three separate decision making tasks; midpoint samples of CMC transcripts and FtF video recordings were extracted from each meeting for coders to rate intimacy measures for each participant. Walther was unable to confirm his hypothesis that initial levels of pro-social behaviour would be higher in FtF than CMC, and then increase to similar levels with time. However, he found that CMC groups rated higher in immediacy/affection, similarity/depth and composure/ relaxation, which are all intimacy-related dimensions of relational communication (Burgoon and Hale 1987). FtF communication was also more task-oriented than CMC for all measurement times and CMC groups became less task-oriented with time. He concluded that "greater task orientation and impersonality associated with CMC previous cues-filtered-out experiments do not occur in extended-time asynchronous CMC interactions" (Walther 1995 p198). The results indicated that when interacting in projects over a period of time, CMC participants adopt a more intimate and sociable relational behaviour than their FtF counterparts from the beginning and throughout (Walther 1995 p198).

Walther's social information processing theory (1992) predicted a difference between CMC and FtF relational communication in terms of rate, not capability (i.e. socio-emotional exchanges happen in CMC at a different rate to in FtF communication, rather than not happening at all in CMC because of media richness deficits). In the 1995 study, however, time dependent interactions were not evident. Based on Gersick's research (1988), he attributes this to the

possibility that such dynamics happen very early in CMC, possibly within the first five exchanges, which his methodology would not have captured. Another interesting area for research might be to compare the time profiles of Tuckman's forming, storming, norming and performing model (see section 2.3.2 under "Development phases") between writing projects supported by varying levels of mediated communications. As a slightly different perspective to the high initial levels of socio-emotional content being attributable to the anticipation of a long term virtual relationship, I believe that it is also possible that CMC has a beneficial effect in accelerating team members to the performing stage. Walther proposes that strategies develop to achieve socio-emotional communication, which therefore increases with time in CMC. However, it is also possible that team members compensate for CMC deficiencies from the start, using more socio-emotional content sooner than in FtF scenarios.

Work by Van der Meij and colleagues (2004) on the interactional coherence in emails between elementary school children, rather confusingly supports Walther's theory and also my own suggestion. 220 emails were sent by 60 groups of children and email content was analysed against a framework comprising contextual, rhetorical and semantic elements. Part of the semantic dimension interpreted the topic as relevant to the communication, personal talk or domain talk. First emails contained 34.8% personal talk, and later emails contained only 4.7%. However, in the study protocol, the children were asked to introduce themselves at the beginning of the projects, so the data may not be representative of natural scenarios. On the other hand, in projects where individuals do not meet FtF, such "introductory" behaviour may be expected early on, rather than later. The fall in personal talk in Van der Meij's study may also be attributable to the participants being children, who are less inhibited than adults to self-disclose from the beginning of a relationship. Van der Meij's results also showed a *decrease* in overall expression of affect between the first and later emails. However, analysing personal, domain and communication talk separately showed decreases in expressions of affect for domain and communication talk, but an increase for personal talk. Frequencies for expressions of affect were 10.0% in personal talk of first emails, and 38.7% in later emails, thus supporting the concept in social information processing theory that socio-emotional exchanges may increase with time.

Walther discusses two possible explanations for the higher pro-social behaviour in CMC in his 1995 study. One explanation was the inclusion of nonverbal cues in the analysis and coder interpretation of video recordings of the FtF communications. Previous studies had not included nonverbal behaviour in analyses and inclusion may have biased coders to interpret non-mediated social behaviour more negatively. Walther's second explanation for pro-social behaviour in CMC was its asynchronous character, which allows people to manage interpersonal exchanges without hindering task activity. CMC appeared to lend itself to self disclosure, whereas in FtF situations, self-disclosure may be interpreted as deviant behaviour, hindering the task which has demanded co-presence and thus removed individuals from other important work. Walther argues, on the other hand, that asynchronous communication awards more individual control over when, how long and how often to participate:

> *...temporal commitments become discretionary, and task versus interpersonal interaction becomes ... de-regulated; both task and social exchange may exist without one constraining the time available for the other... CMC provides an 'electronic water cooler', where employees may both do 'Job talk' and 'shoot the breeze' conveniently, without having to leave their desks and without risking the impression that they are not working (Walther 1995 p199).*

This last statement brings me back to my emphasis on the difference between *opportunities* to communicate, which encourage communications and therefore collaborations between co-located researchers (Kraut et al. 1988), and the quality of the communication channel. There is no virtual 'water cooler' for chance meetings in a virtual team communicating by email, but perhaps the added value in terms of individual control with asynchronous email communication encourages wilful informal exchanges, thus compensating for the lack of chance opportunities presented in FtF scenarios.

Similar to Walther, Chidambaram (1996) also criticizes research which has only focused on single session studies. Such research does not allow for the fact that although computer support may initially lower relational intimacy, teams may eventually develop ways of exchanging socio-emotional communications. In an experiment with 28 five-member groups, half with group support systems (GSS) and half without, four decision-making tasks were completed in four separate meetings. Measures collected at each meeting were cohesiveness, perceptions of process and satisfaction with outcome. These dependent variables were affected by time and treatment (GSS/non-GSS). Attitudes of GSS users changed over time from highly negative to somewhat positive and outcomes improved more slowly. The changes in attitudes of the groups over the four periods challenge the notion of constancy in media effects, with some evidence of relational affiliation over time among groups using GSS. Thus repeated use of the system over time increased the group's affiliation, providing support to the social information processing theory.

However, findings from later research by Burke and Chidambaram (1999) partly challenge the concept that relational communication supporting social presence may become equivalent between different media with time. In their experiment comparing FtF, synchronous and asynchronous communications over an ambiguous task, perceptions of social presence after the first and fourth meetings over a four week period remained constant, with the FtF group having significantly higher perceived social presence than the two mediated conditions. The same remained true after the last meeting. (Interestingly, perceptions of social presence did not vary with synchronous and asynchronous conditions as expected based on the premise that immediacy in response might improve feelings of social presence.) On the other hand, perceptions of communication effectiveness differed significantly between the FtF and synchronous groups at the beginning, but this difference became insignificant by the end of the research, thus partially supporting the time-based aspect of the social information processing theory.

Adaptive structuration, a concept originally formulated by a British sociologist, Anthony Giddens, acknowledges the reciprocal impact people and systems have on each other (Burke and Aytes

1998), in that technological (bandwidth etc.) and social structures (rules and norms) constrain and enable group activities. In support of both adaptive structuration and social information processing theory, Burke and Aytes studied different types of media, FtF, video conferencing, synchronous and asynchronous CMC etc. in a group writing scenario with 238 participants across 62 groups. This was a longitudinal study across four sessions. Even though tasks in this study were highly equivocal, richer media did not result in higher satisfaction. The results showed that group cohesion and performance increased over time for all media, irrespective of richness variations, reaching roughly equivalent levels by the final sessions (Burke and Aytes 1998). Although an experimental study with student subjects, the findings provide strong evidence for these two more social interactive theories of mediated communication.

Review of the theories of mediated communication

The theories of mediated communication which I have discussed can be reduced to two schools of thought, deficit theories and the more social interactive theories.

The deficit theories, (media richness, lack of social context cues and social presence theories), are underpinned by technological attributes of the media, suggesting that certain media are restricted to task oriented exchanges. Here lies a paradox: the social dimension of team working contributes positively to performance (see section 2.3.2), but CMC supporting distributed team work, as described by the deficit theories, appears to promote task oriented communication and inhibit socio-emotional communication, thus precluding the development of team cohesion. Indeed, this is confirmed by Pauleen and Yoong's participatory action research study of seven organizational professionals facilitating virtual teams. Their respondents interpreted email as a "channel more suitable for communicating information and coordinating projects than for building relationships" (Pauleen and Yoong 2001 p199). Deficit theories suggest that the social dimension of networked team work may be disadvantaged.

On the other hand, the more social interactive theories, (social influence, deindividuation, social information processing and adaptive structuration theories), begin to take context, relationships and dynamics into account, suggesting that interactions will develop strategies to equalize levels of socio-emotional exchanges in CMC to those in non-mediated communications.

Social influence theory offers a model of mediated communication more in line with Nystrand's social interactive theory of writing, where media usefulness becomes a social construct, rather than a technological given. Deindividuation introduces a development of social norms and the social information processing theory takes a step once again towards a social interactive stance and adaptive structuration, suggesting that with time, socio-emotional strategies in exchanges will develop. Indeed, many strategies of socio-emotional communication in emails have been researched, and I now move on to discuss some of these.

2.3.5 Socio-emotional mediated communication

Overview

> *"Ethically valuable social relations, a subset of social relations generally, are relations in which all those involved flourish" (Rooksby 2002 p2)*

As I have shown in the previous section, there is much academic debate over whether the leanness of media affects the capability of individuals regarding socio-emotional communication. There is also much research specifically focused on socio-emotional content in CMC, reinforcing the concept that it can convey other than purely task-oriented information. Socio-emotional communication in emails can be represented by paralanguage, politeness strategies, self-disclosure, expressions of affect, and explicit expressions of presence, and I discuss some research on the use of these strategies here.

Theories and research

Lea and Spears (1992) have investigated paralanguage as a means of communicating social information. They analysed paralinguistic cues in emails from discussions among 16 groups of three students. Half of the groups were told individual styles of communication were being studied and the other half were told that the group style was under focus. This created high and low group salience conditions. Further, members of two of the groups were seated in the same room, but told they could only communicate via computer, whereas for the other two groups, members were seated in separate rooms to create visual anonymity. This created individuation and de-individuation (visual anonymity) conditions respectively. For the de-individuated group with high salience, paralanguage use correlated positively with person perception, whereas it correlated negatively in the group with low salience. Thus when subjects were de-individuated and group salience was high, the meaning placed on the paralinguistic marks conformed to the social attraction response associated with subjects' sense of group identity. When group identity was low and individualism the salient context, paralanguage use was interpreted negatively. For group members seated together, the effect of a group context on the perception of paralinguistic cues was reduced. These were interesting conditions to study, because the benefits the SIDE (social identity and de-individuation) model bring to virtual teams are affected by the social (imposed sense of group) and environmental (same room or apart) conditions imposed. These conditions thus mirror real life scenarios, where groups may be established within (high group salience) or across departments (lower group salience) and some team members may be co-located and others remote. Although only based on an experimental study, Lea and Spears' study shows how a communication strategy, use and interpretation of paralanguage, are adapted according to the social contextual use and contribute towards a positive social identity in the scenario of team work using email.

Self-disclosure and expressions of positive and negative affect represent uninhibited behaviour, and are further elements of socio-emotional communication. Sproull and Kiesler (1986) discuss the positive influences of uninhibited behaviour on the task dimension, quoting communication

and innovation literature. The structural and social barriers which impede communication and innovation are removed with email, allowing creativity, and leading to new ideas (Sproull and Kiesler 1986 p1511). However, there has been some research to suggest that both task and affective conflict are detrimental to performance in distributed teams (Mortensen and Hinds 2001 p231) and certainly negative interpersonal behaviour in emails may jeopardise effectively building the social dimension of a virtual team. In Sproull and Kiesler's field study results, flaming (negative affective interpersonal communications) was reported in emails 33 times a month, and in FtF conversations 4 times a month, and one of the researchers' conclusions was that "people behaved irresponsibly more often [in email] than they did in FtF conversations" (Sproull and Kiesler 1986 p1509). Results of more recent research by Mortensen and Hind (2001) and Jacqueline Taylor (2000), however, do not support Sproull and Kiesler's negative valence of flaming associated with email communication.

Mortensen and Hinds (2001) completed a survey of 24 teams from 5 companies, 50% co-located, 29% domestically distributed and 21% internationally distributed. These researchers analysed affective conflict, task conflict, shared identity and team performance and did not find more affective conflict in distributed teams. Taylor (2000) found in her study of email discussions set in a working environment, that flaming was highest in an individuated, more identifiable condition, than in the impersonal, more anonymous scenario. Conditions were two levels of anonymity, provision of group member details or names only, and two levels of group salience, manipulated through the instructions given. There was also more self-disclosure in the less anonymous scenario, i.e. the more subjects knew about others, the more they tended to disclose about themselves. Frequencies of messages were higher in the less anonymous groups although communications were more evenly distributed between individuals in the more anonymous groups. Analysing measures of interpersonal perception, Taylor found that group cohesion was highest for groups with higher identification, and that group salience did not affect group cohesion. Thus with higher frequencies of communication, flaming, self-disclosure, and group cohesion in the individuated group, Taylor's results argue against the technologically determined effects of reduced social context cues, and rather than deindividuation, individuation appears to have promoted uninhibited behaviour. Providing identity information in this research promoted socio-emotional communication in terms of flaming and self-disclosure and promoted group cohesion (Taylor 2000 p194). Taylor's findings thus conflict with the concept that anonymity encourages self disclosure and expressions of affect, and suggests instead that an intervention (making information about team members available) is a strategy to promote socio-emotional communication and encourage group cohesion.

Tidwell and Walther's experimental study in 2002 with 158 student participants compared CMC and FtF interaction in first meetings. The CMC setting was a semi-synchronous system based on an email system; messages were transmitted to remote partners on completion. Findings showed that CMC users adapt to the medium through the modification of uncertainty reduction behaviours. Without nonverbal cues, CMC partners abandoned the socially acceptable questions and answers characteristic among new acquaintances in FtF situations. CMC

participants adopted "more direct, interactive uncertainty reduction strategies – intermediate questioning and disclosing with their partners – than did their FTF counterparts. The probes and replies they exchanged were more intimate and led to levels of attributional confidence similar to their offline counterparts'" (Tidwell and Walther 2002 p339). Tidwell and Walther discuss the paradox that personal questions and self-disclosures "offering potentially individuating information, reinforce the presence of social, and the lack of individual identity" (Tidwell and Walther 2002 p340). In this study, similar to Taylor's (2000) research discussed above, email did not appear to inhibit self-disclosure. Tidwell and Walther's work also supports the concept of adaptive structuration playing a role in email communications, in which interactions, rather than technology or individual attributes, are influential.

Tidwell and Walther's work suggesting that less inhibited behaviour may be a strategy interactively developed between individuals and CMC to achieve socio-emotional exchange and social identity and Taylor's work showing that providing personal information about team members increases self-disclosure and expression of affect still leave a dilemma; strategies which promote socio-emotional exchanges may influence both the negative and positive expressions of affect. In Taylor's work, there was increased flaming with the intervention. Increasing both positive and negative affect may have conflicting influences on team solidarity and performance, a dilemma emphasizing the importance of positive politeness strategies in email communications.

Politeness strategies are another example of how email communications can present a positive valence contributing to team solidarity. Sandra Harrison (2000b) has applied Brown and Levinson's (1987) framework of politeness strategies in spoken discourse, to an analysis of politeness strategies in 23 consecutive emails from a naturally occurring email discussion group. Harrison found many instances of politeness strategies in the email discourse she studied. Of particular note in this study was the fact that participants were using predominantly positive strategies, which reduce social distance and relative power, thus promoting discussion in a safe atmosphere, and strengthening the group (Harrison 2000b p78).

As a final aside to socio-emotional communication, but still focusing on social presence built through email communication, recent work by Panteli has focused on explicit articulations of presence, rather than socio-emotional representations. Panteli's definition of social presence is the state of being "there" (Panteli 2004 p73). Panteli argues that mediated language "not only reflects a specific virtual context, but it also helps in its production, reproduction and transformation" (Panteli 2004 p62), and uses discourse analysis, therefore, to "unpack the creation and ongoing recreation of patterned social relationships" (Panteli 2004 p62). He collected 432 emails from a real-life virtual project involving 25 remote team members. Categorising main themes of emails as related to the forming or performing phase, Panteli identified how participants talked about their own and others' presence in the virtual team environment. In such articulations or absence of such articulations, writers are thus providing a form of self-disclosure and informing on their availability.

Articulations about time and on line availability reflected presence in the forming stage, and extended in the performing stage to include references to commitments in other "contexts" i.e. to other work assignments and to personal matters. "Presence was discursively negotiated and renegotiated and constructed even through words and emails that were never said" (Panteli 2004 p73). He gives the example of writers talking about their absences, which implies a "do not disturb" message. "The negotiations that they enter into with their team members to define their presence in the shared-mediated environment have contributed to forming and maintaining boundaries between ... environments" (Panteli 2004 p75). Panteli thus contributes to the debate over whether boundaries to virtual teams exist, by showing that individuals actively create boundaries between their shared mediated and non-shared environments through the messages they articulate, and also through implied messages articulated through silence.

This aspect is of particular significance to socio-emotional exchange in networked teams. In real or virtual teams, Panteli points out that "members are expected to be present and to develop personally engaging behaviours in role performance" (Panteli 2004 p77). "The virtual context, by its nature allows team members to be 'absent' and 'silent', which can contribute to relationship problems and feelings of isolation". Thus Panteli highlights a separate issue to that of the opportunity of transmitting rich unambiguous information as afforded in FtF scenarios, and that is lack of social restrictions or boundaries to the virtual team construct, which affords the *opportunity of absence*, not afforded in teams which are physically co-located. This is an interesting issue similar to the first topic I discussed in this section on networked teams, on the communication opportunities offered by physical presence – in this case we are discussing the absence opportunities offered by virtual presence.

Review of socio-emotional mediated communication strategies

In this section, I have discussed a number of socio-emotional exchange strategies which can be used with CMC. Paralanguage use supports socio-emotional exchange and social identity (Lea and Spears 1992); the intervention of providing personal information on team members encourages self-disclosure and group cohesion (Taylor 2000); self modification of uncertainty reduction behaviours results in less inhibited behaviour, such as increased self-disclosure (Tidwell and Walther 2002); positive politeness strategies reduce social distance, contributing to team solidarity (Harrison 2000b), and explicit articulations on absence and presence provide an element of self-disclosure (Panteli 2004) and can inform social boundaries.

Thus research has shown not only that email can support socio-emotional exchanges, but that specific strategies have developed to achieve this, in support of the social interactive theories of mediated communication.

2.3.6 Email communication

Overview

Bearing in mind the value of the social dimension to team work, and that in networked team writing (the focus of my research) communication is mediated by email, I have discussed above whether email encourages or discourages socio-emotional exchanges. Drawing from research into socio-emotional exchanges in CMC, I concluded that email can support the social dimension in team working. I now turn to discuss whether team behaviour is identifiable from email communication.

Later in sections 2.4 and 2.5, I argue the case for developing an email analysis tool for researching networked team writing projects. This next part of my literature review aims to show how email style may inform networked team writing, thus to introduce the concept of analysing project *email communications* for *team writing research*. I therefore discuss style and metadiscourse in emails, and then draw on some examples of research which has interpreted social behaviour from emails.

Email style and metadiscourse

> *...textual style is a subset of a person's style in general...style is best seen as those qualities of people's performances of social practices that express their attitudes, interests and character, in short their selves, to other people (Rooksby 2002 p10).*

Many researchers are debating whether the language of CMC veers towards written or spoken discourse, or has perhaps developed a style of its own (e.g. Baron 2001; Ferrara et al. 1991; Harrison 2000a; Yates 1996; Yates and Orlikowski 1993). Rooksby too, (quoted above) recognizes other influences on textual style than the self alone: "But textual style, like the styles of art objects more generally, may also be taken to consist of formal properties attaching to artefacts, without any reference to the performances of those artefacts' creators" (Rooksby 2002 p10). She points out that any study of style in CMC should consider the relations between the text and writer, text and reader and between the text and the world (Rooksby 2002 p11). Mirroring Nystrand's (1989) social interactive theory of writing, she argues that "style covers both the performance and the interpretation of significant objects and actions, and cannot be determined by either producer or a receiver alone" (Rooksby 2002 p15). Neither the writer nor the reader therefore has complete control over what is stylistically significant, and the style enacted by individual performances situated within social practices cannot be accurately defined by either.

Danet (2001a) analyses emails with a view to evaluating the development of stylistic norms in exchanges. She concludes that email technology invites informality, although the informality is not necessarily attributable to the technology alone. She suggests that such a trend could also be attributable to a historical shift in genres in English, in both personal and professional contexts. She describes a move away from the traditional view that oral discourse needs to

follow the rules of written communication to be intelligible; rather the view has emerged that written communication needs to adopt the style of oral discourse to be intelligible. Danet (2001a) attributes this trend as partly due to the Plain Language movement in the US and UK in the 1970s, which in the interest of clarity in written communications naturally encouraged an active and more personal style of writing, more similar to oral discourse than the earlier style of bureaucratic language and 'legalese'. Other drivers of this trend are the changes in approaches to the teaching of writing, and the shift towards a postmodern, more active style of writing in academia in the 80s and 90s. Finally, she points out that what we bring individually to our email writing affects our style. The generations who have grown up with less formal written styles in an established world of mediated communications generally approach email writing in a different way to those of us whose education preceded the Internet uptake, and who are therefore more likely to be entrenched in positivist objective writing styles. Danet concludes that "the language of email was in a state of transition as we approached the millennium" (Danet 2001a p93). She predicts that email style will become increasingly less formal, particularly regarding greetings, and that the differences between official and personal emails will lessen; this style will become expected and therefore accepted as legitimate; variation in public-official email writing will be greater than in traditional letter writing, but certain letter-writing characteristics will persist in certain sensitive scenarios, particularly in first emails which represent a virtual first meeting, in upward communications to people of higher status, to strangers, and where there is high risk for the writer. She also predicts that "as email matures... different text-types will come to have different degrees of normatively approved formality" (Danet 2001a p94), with the more normatively formal style paralleling formal text-types on paper. Her final prediction is that younger people, unaccustomed to traditional letter-writing, will adopt the new style, even in the sensitive scenarios described above. "They will do so with little ambivalence or uncertainty, and will feel comfortable introducing playful material, e.g., a signature file, even when the rest of the letter is in a serious frame" (Danet 2001a p94).

In their study of 280 emails, Goldstein and Sabin (2006) coded email acts and identified related genres. Having identified these they tested the classification and achieved reasonable performance for five email act categories and two genres. They believe their "findings support the characterization of email as an amalgam of unique communicative genres, where the common genre – email conversations is most similar to spoken communication" (Goldstein and Sabin 2006 p7).

Postmes, Spears and Lea (2000) studied the evolution of communicative norms in emails amongst students. They showed that content and form of communication is normative and defined by group norms, conformity to group norms increases over time and communication outside the group has different social norms. They argue that "the content of communication within CMC will be contextually determined and influenced not only by the general norms of the subcultural milieu (e.g., McCormick and McCormick, 1992), but also the specific local norms and practices of the communicating group" (Postmes et al. 2000 p366). They conclude from

their study that the content and form of messages are variable, socially structured, and subject to emergent norms specific to one's social group.

The concept that forms of CMC may develop their own hybrid genres, registers or styles is supported by many researchers. Ferrara et al. (1991) identify real time interactive written discourse as an emergent register, and Baron analyses email in the framework of a contact language to understand its schizophrenic (part speech, part writing) nature (Baron 2001 p258). She also predicts the likelihood of two styles of email, one formal (edited) and one informal (unedited). She points out that frequent email users may switch off automated editors and may even choose not to manually edit, thus communicating in an informal way, whereas on the other hand "a contract is still a contract" (Baron 2001 p242), requiring accuracy, editing and thereby more formal communication. Crystal (2001) too recognizes the increasing use of email in professional settings in addition to its use for more informal personal communications.

> *The result will be a medium which will portray a wide range of stylistic expressiveness, from formal to informal, just as other mediums have come to do, and where the pressure on users will be to display stylistic consistency in the same way that this is required in other forms of writing (Crystal 2001 p128).*

Gains (1999) throws a different perspective on the question of whether email tends towards written or oral language, showing a distinction between academic and commercial discourse communities. In a small scale study of 116 emails in academic and commercial settings he found that academic emails were less formal with more social chat, i.e. more like conversation, whereas emails from a commercial environment were more like written business language.

From an empirical perspective, Yates (1996) has shown that speech, writing and CMC can be differentiated through type/token, lexical density, pronoun use and modality analyses, thus identifying characteristics which define these genres. CMC and writing were similar in terms of type/token ratios and lexical densities; pronoun and modal auxiliary use appeared to be more similar between CMC and speech, however. Yates and Orlikowski (1993) also argue from their quantitative analysis of 1353 messages between computer language designers using the ARPANET in the early 80's, that email shows "characteristics of both written and spoken discourse, as well as characteristics seemingly unique to electronic discourse" (Yates and Orlikowski 1993 p13).

To close on this debate over whether email has its own style(s), the influences on that style and whether such a style is yet fully developed, the important point relevant to my research is that there are identifiable traits in email styles, that these traits can be quantified, and that by analysing them, we may be able to explore deeper to understand their significance, influence on and representation of interpersonal behaviour and relationships in team work. Researchers of writing and communications support this concept: Faigley writes that "Words carry the contexts in which they have been used" (Faigley 1985 p240). Yates writes "Not only must the text carry the social situation, it must also carry the participants' relationship to the situation, their perception of the relationships between the knowledge and objects under discussion" (Yates

1996 p46). With these elements carried in an email text, we should be able to extract and interpret their representations.

However, all is not so simple; researchers recognize the confounding effect of CMC being both "the source of norms and the place to observe them at work" (Gains 1999 p346). The study of communication to understand behaviour, when the communication shapes that behaviour and the behaviour shapes the communication, presents a methodological dilemma, particularly where the communication is the focus of all social interactions. I visualize this problem as though communication and behaviour are on either side of a sheet of very clear glass. With a virtual team, the glass between the communication and behaviour is so clear, that I am unsure whether it exists; communication may *represent* or *be* the behaviour. However, I turn now from the abstract to some concrete examples of research which has studied email communications and interpreted social behaviours.

Vaes et al. (2002) have demonstrated a surprisingly simple marker of pro-social behaviour in emails through the use of first person pronouns. The researchers explored whether "in an interpersonal context, mentioning oneself is to become more involved in the situation and to increasingly relate oneself to the other person" (Vaes et al. 2002 p 527). They tested whether the use of person pronouns varied when the relationship between the sender and the addressee became more intimate. They asked one group of participants to imagine they had received a "lost" (i.e. incorrectly addressed) email from a stranger. A second group were asked to imagine they had received a message from a friend who had mistakenly used their address. Participants were asked to write a reply as though it was a real life situation. Respondents to friends used significantly more first person singular pronouns than those who answered to a stranger. One limitation to this otherwise extremely robust study is that it was limited to the French language. However, Brown and Levinson's (1987) conclusions on the commonality of linguistic strategies for politeness across languages suggest that such findings may be applicable to other languages, and indeed pronouns are used elsewhere to research communication strategies and style (Eggins and Martin 1997; Hyland 1998; Te'eni et al. 2001; Yates 1996).

An example of a social influence being extracted from email metadiscourse is given by Sherblom's early work in 1986. His research showed a difference in signature behaviour with organizational hierarchy, which interestingly conflicts with the concept of status equalization suggested by Sproull and Kiesler's work (1986). He writes "an electronic paralanguage reflects reinforces and recontextualizes the organizational structural hierarchy" (Sherblom 1988 p50). He thus suggests that the style of communication in email reflects the power relationship between the correspondents. In a small scale study of my own analysing 293 emails from a professional writing project, I also found variations in greeting length according to the level of the recipient in a hierarchical organization, and variations in word count according to the direction of email transmission in the hierarchy (Edwards et al. 2005 p756; Edwards et al. 2006 p181).

Nickerson (2000) discusses typical communicative practices identified in a representative sample of English emails written by British and Dutch speaking employees in a large multinational corporation. She collected emails written over a two month period from seven managers and analysed a sample of 100 English emails written by Dutch speakers and 100 English emails by (British) English speakers. She discusses similarities between the strategies of Dutch and British writers and the possible cultural and organizational influences. She found a "certain amount of relational or non-propositional content, intended to maintain the social system with the corporation, i.e. the patterns of corporate social relations between employees" 2000 Nickerson p153). She observed from her analyses that "salutation was only included if the message was sent to a single primary receiver and if it was included it always took the form of the first name of the recipient" (Nickerson 2000 p156). Messages with more than one recipient had no open greetings. Nickerson also reported a presence or absence of a pre-close statement, such as "Looking forward to hearing from you". The final close in the corporate emails was pre-programmed and varied in language use and formality. Nickerson concluded:

> *...code used in the text of the message had little effect on the code used in the [close], as might be expected in other forms of written business communication, such as a business letter, and the use of either or both languages was viewed as an appropriate formal realisation by those members of the corporation represented in the data set (Nickerson 2000 p157).*

Nickerson discusses the possibility that such pre-close statements are markers of politeness, possibly used where the corporate distance between correspondents is larger (interdivisional rather than interdepartmental) or when a degree of compliance is required from the receiver.

Nickerson investigated several interpersonal markers, two of which were politeness strategies and first person pronouns as markers of involvement and solidarity.

> *The total number of occurrences for each pronoun suggested some divergence between the two groups of writers [Dutch English writers and British English writers] in their use of 'I' and 'we' respectively. The Dutch writers showed a preference for 'we' as a first person pronoun, together with a preference for the use of 'we' as a pronoun excluding the receiver [74% exclusive], whereas the British writers showed a preference for 'I' and an inclusive use of 'we' [24% exclusive], including both sender and receiver (Nickerson 2000 p173).*

For markers of politeness, there were some similarities and differences between the two writer groups. Requests were nearly always modified by the inclusion of politeness markers e.g. "I would appreciate it if.." Only British writers used the expression "Perhaps, maybe you could/ would/should be so kind as to..." The most commonly used expression by both Dutch and British writers was "Please...". Nickerson concluded agreement with Mulholland (1999 p81-81) that writers use politeness strategies in emails regardless of the preference for minimalism.

Overall Nickerson concludes from her email research that "a typified corporate discourse may exist regardless of the national culture of the individual employee" (Nickerson 2000 p176). Thus through the study of email metadiscourse, Nickerson was able to draw conclusions regarding the social practices in the particular situational context under focus.

Review of email research

In this section I have drawn on the views of experts and from examples of research to show that email style combines characteristics of less formal spoken language, and also of more formal traditional business letter writing. There appears to be some consensus that email style is still in evolution and that styles will diverge according to purpose, for personal or professional communications. Finally using examples of work from Vaes (2002), Nickerson (2000) and Sherblom (1988), I have shown how the style of email, extracted from interpersonal markers such as pronoun use, greetings, signatures and politeness strategies, can be used to interpret and understand situational social behaviour.

2.3.7 Relevance of CMC and team theory to this research

There is strong evidence from the literature (Evans and Dion 1991; Kelly and Duran 1985; Mortensen and Hinds 2001 p231; see under section 2.3.2) to suggest that the social dimension contributes positively to team work. Opportunities for socio-emotional exchanges in team work are therefore important, and some early research (Kraut et al. 1988) suggests that the quality of *unplanned* informal exchanges in co-located teams is especially significant. Team behaviour is dynamic in nature, and to develop and maintain a social dimension in networked team writing requires socio-emotional exchanges by email. However, media richness, lack of social context cues, and social presence theories explain CMC in relation to its deficits compared to non-mediated communication, and suggest that email exchanges may not support socio-emotional communication as effectively as FtF exchanges.

Theories of CMC have broadened in the same way that writing research has evolved to adopt a more social interactive perspective. Social influence, deindividuation, social information processing and adaptive structuration theories of CMC view its use from the interactive and situational perspective. These theories suggest that email may offer benefits over FtF scenarios, in terms of media perception and use (social influence) increased group cohesion (deindividuation), and interactive adaptation of technology and individuals to achieve similar levels of socio-emotional exchange as those possible in FtF situations (social information processing and adaptive structuration). Indeed, research has shown that email *can* support socio-emotional communication, that certain strategies are used for this purpose in a positive way (Harrison 2000b; Mortensen and Hinds 2001; Sproull and Kiesler 1986; Taylor 2000), and that these strategies can promote a sense of team belonging and group cohesion (Harrison 2000b; Sproull and Kiesler 1986; Taylor 2000), which benefits team performance.

Finally, I touched on the ongoing debate over the style of email, and whether it is more similar to spoken or written language. Email style appears to be diverging into two styles, one more formal for professional purposes and another less formal for personal purposes. Using examples of work from Vaes (2002), Nickerson (2000) and Sherblom (1988), I have shown how the style of email, extracted from interpersonal markers such as pronoun use, greetings, signatures and politeness strategies, can be used to interpret and understand situational social behaviour.

Based on the concept that strategies are used in email to support socio-emotional communication, which improves group cohesion and thereby team performance, this research compares socio-emotional communication behaviour and project performance for two networked writing teams. I use interpersonal markers from emails to interpret adaptations in socio-emotional communication behaviour for the audience, context and purpose. Additionally, I develop a formality scale based on the concept that email style varies between that of traditional business letters and spoken language. My research analyses email behaviour against Nystrand's social interactive model of writing, for which I explain the rationale in the following section. The methods of extracting interpersonal markers are explained in section 3.7. Chapters 5 and 6 report case studies on the communication behaviour of two separate team writing projects, and chapter 7 compares the results and performance of these two projects. Finally in chapter 8, I return to review the findings from my research in the light of theories presented here.

2.4 Mapping email style against the social interactive model

In Nystrand's model of writing, the text is not autonomous, but a communicative event with a context of production and reception (Nystrand 1989 p73). The text is the negotiation of meaning between writer and reader. "We conceptualize text meaning, not in terms of the writer alone, but in terms of interaction between writer and reader purpose" (Nystrand 1989 p76). Accepting Nystrand's social-interactive model of writing, and writing as a communicative event raises the question of whether communication competencies during the process of team writing, aiming to maintain the group (i.e. communication and coordination activities throughout the process) might not reflect the communication competencies required in the team writing of the document (Edwards 2005b; Edwards et al. 2006 p174). Interpersonal skills required to maintain the group are reflected through the team members' abilities to achieve shared understanding in their team communications, which in networked teams, are often mediated through emails. If models of writing processes apply equally to written emails as they do to the final documentation in writing teams, we might expect influences on writing processes to influence communication behaviours in emails. The first hypothesis explored in this research, therefore, is whether writing influences affect communication behaviour in team writing projects and whether content analysis of emails alone can deliver both of sets of variables:

> *Pilot study:*
>
> *H1 = Email communication behaviour is the product of writing influences and representative variables of both can be derived non-intrusively from email content.*

Choice of semantics, rhetoric and exchange patterns in email communication provide a hidden discourse. The measurable data in email analysis are numerous, including choice to use email, exchange patterns, frequency, length, speed of response, and the diverse aspects of actual content analysis. (Examples of research are discussed in section 2.3.6.) These measurable data provide indicators of social constructs, for example such as formality of relationships

between members, and behavioural predictors and have both exogenic, quantitative and endogenic, qualitative characteristics, thus meeting the multiple realities and mixed methodologies epistemological stance discussed in section 2.2.6.

> *The most important contribution of discourse analysis is that it provides a way to unpack the production of social reality...Discourse analysis provides the tools to understand the social processes that produce organizations. This contribution is even more important when we consider that many of the more recent topics that have been the focus of intense research activity, such as the knowledge-based firm, the virtual organization....We require new approaches if we are to understand the dynamics of these new phenomena (Phillips and Hardy 2002 p82).*
>
> *The unique contribution of discourse analysis...is to insert the discursive level to understand how structured sets of text and the practices of their production, dissemination, and reception, together constitute the social (Phillips and Hardy 2002 p86).*

In this research, I study the email texts exchanged in an attempt to unpack the social reality throughout a team writing project, addressing the following hypothesis:

> *H2 = Social dimensions of teams can be identified from email communications.*

Further, in section 2.3.2, I justified how the social dimension of team working contributes to performance. Knowing that pro-social behaviour promotes group cohesion and that group cohesion promotes team performance, we might expect writing teams showing more pro-social behaviour to perform better than writing teams showing less pro-social behaviour. Based also on Nystrand's social interaction theory of writing processes, adaptations of communication behaviour (in emails or other documents) in anticipation of readers' needs, increase the chances of a shared social reality of meaning between the writer and reader, and effective documentation. Hyland suggests that research into metadiscourse "may reveal that the ways writers control the expression of textual and interpersonal relationships within a text are as vital to the rhetorical success of a text as its propositional content" (Hyland 1998 p241). With team activities aimed at meeting socio-emotional and task needs, email communications on networked team writing projects should therefore reflect the balance between the task and socio-emotional dimensions, helping to predict the performance of the final project deliverable:

> *H3 = Social interactive adaptations and pro-social behaviour in a writing team's emails are reflected in the relational metadiscourse and social desirability of the document produced by the team.*

I now summarize the rationales developed from this literature review, which substantiate researching both the development of an email analysis tool and the concept that social interactivity and pro-social behaviour in emails reflect social interactivity and social desirability of the final document in team writing projects.

2.5 Rationales for my research

My research attempts to answer the question: "Can we learn about the influence of team culture on virtual team writing from content analysis of email communications during projects?" The three hypotheses used to answer this question are:

> *Pilot study:*
>
> *H1 = Email communication behaviour is the product of writing influences and representative variables of both can be derived non-intrusively from email content.*
>
> *H2 = Social dimensions of teams can be identified from email communications.*
>
> *H3 = Social interactive adaptations and pro-social behaviour in a writing team's emails are reflected in the relational metadiscourse and social desirability of the document produced by the team.*

Different positions on knowledge, such as positivist, quantitative and scientific vs. interpretive and qualitative, influence how research questions are defined, choices of methodology and the knowledge acquired. To help validate my research, therefore, I adopt the stance of multiple realities of knowledge – that some is in the mind, some is outside of our minds and can be found, and that some is created within and among minds – and adopt both qualitative and quantitative approaches. This research uses endogenic, interpretive and exogenic, positivist methods to support each other and avoid the bias of adopting a single stance.

This research adds value to team writing research by combining different research perspectives, textual, individual, group and social, to study current, real virtual team writing practices, and by developing a standard research methodology for cross context comparisons of case study findings. Standard data collection and analysis techniques allow results from different studies to be compared and meta-analysis of data from multiple contexts, from which research findings may be more broadly applicable.

Additionally, the method includes the study of process and outcome towards developing a causal model, to understand how team culture can influence the outcomes of virtual team writing. Input variables are writing influences; these together with task and group maintenance activities during the process are profiled from email communications and analysed against Nystrand's social interactive model of writing; output variables are derived from participant feedback on the projects and document evaluations. Data is searched for causal relationships between the writing influences, communication behaviours and quality of the resulting process and document, to identify any hidden constructs or combinations of variables, which influence the team culture and the end product.

With application of the same communication behaviour to written email communications and to the end product of team writing projects, content analysis of written emails provides a proxy method of research into professional writing practices. Analysis of email records retrospectively on writing projects provides an accurate, non-intrusive technique to study writing processes, with no researcher intervention to the actual process and no additional tasks introduced in the

work context. Email analysis thus provides an ideal tool for the study of professional writing. My research therefore includes the design of such a tool to identify social constructs and group evolutions in writing research.

The originality of this research lies in the fact that it analyses project *email communications* for *team writing research*, an approach which has not previously been reported in the literature.

3. Methods: towards an email analysis tool

3.1 Chapter overview

In this chapter, I first discuss ethical considerations and the research design in sections 3.2 and 3.3. In section 3.4, I describe my search for data and provide background to the three projects chosen. The case studies reported in this thesis use four sources of data from virtual writing projects: pre and post project interviews with key participants, records of emails exchanged between collaborators during the project, and questionnaires completed by participants. Methods were applied as appropriate for each context and for the aspect of the research framework under focus for each case study (see Figure 1-1). I explain the methods used in sections 3.5 to 3.8. In section 3.9, I discuss reliability and validity issues. The results of the case studies are reported in chapters 4, 5 and 6. This chapter serves the primary purpose of describing the methods of data collection and analysis used, and also serves a secondary purpose of showing how I improved the methods across the three studies, towards developing a standard email analysis tool.

3.2 Ethical considerations

This research involves coding the content of interpersonal emails. Example codes are for opening or closing greetings and for task or socially-oriented content. Subject matter is not relevant to the research and identities of email authors are not disclosed. Questionnaires or interviews are designed based on the results of the email analysis, to validate interpretations of team work activity in two of the projects. Guidelines drawn up by the European Science Foundation (2000) to promote ethics and good practice standards in science and improve credibility of results were studied for the design of this research, together with relevant literature on research ethics from Sheffield Hallam University (SHU) and the Association of Internet Researchers (AoIR).

SHU Research Ethics Policies and procedures (SHU 2005) list four relevant criteria:

1. Beneficence and Non-malfeasance. There are no risks or harm anticipated to original authors of emails through this research. The research does not involve sensitive topics likely to cause significant embarrassment or discomfort to participants, or relating to highly personal information, or to illegal activity.
2. Informed consent. There are no active "participants" in this research. Email communications from projects are analysed retrospectively, with the written informed consent of representatives of the organizations in which the projects were completed, and who have provided the email communications for analysis.

3. Confidentiality/ anonymity. Names of original email authors are not reported to protect anonymity and confidentiality. Company confidential information and intellectual property are respected and are not disclosed in this thesis.
4. Authority. Following consultation with the SHU ethics committee, and with my supervisors, and in view of the lack of risk or harm to original authors of emails, this research was completed with written consent of at least one authorized responsible from each project.

Further in consideration of AoIR recommendations (Ess and AoIR 2002 p7-8), the "subjects" in this research are authors, subjected to no type of medical or other research intervention. AoIR recommendations acknowledge the lower risk in research which addresses form rather than the content (Ess and AoIR 2002 p7-8), which Danet, a member of the AoIR ethics working committee, also discusses:

> *... if only aspects of linguistic and textual form are being studied, and it is not possible to obtain consent of one or both parties, publishing letters in whole or in part is not a violation of the rights of human subjects, providing that identities are disguised (Danet 2001b p33).*

Having considered the sources cited above and consulted with experts at SHU, I conclude that the research poses no risk to participants and is ethically sound.

3.3 Research design: case studies and email content analysis

There are a number of reasons why case study research is particularly appropriate for researching writing and why the findings may be applicable to a broader field. I discuss these points here, together with the rationale for using email content analysis to research professional writing.

Case studies focus on real contexts. If we separate research from the true professional practice of what we are researching to understand the skills required for that practice, we not only remove the situatedness in which that practice exists and to which it relates, but we distort the reflected practice by imposing an artificial situatedness irrelevant to the practice in its true context. Flyvbjerg writes "in the study of human affairs, there appears to exist only context-dependent knowledge" (Flyvbjerg 2006 p221) and "social science has not succeeded in producing general context-independent theory, and thus has in the final instance nothing else to offer than concrete context-dependent knowledge" (Flyvbjerg 2006 p223). He also argues that generalizability of case studies can be increased by the strategic selection of cases (Flyvbjerg 2006 p229) and that they contain greater bias than other methods toward falsification rather than verification of researcher's preconceived notions (Flyvbjerg 2006 p237). Where case studies research situated actions in a consistent way, the findings can be compared towards a wider understanding and application. In his discussion of the current debate around case study research in action research, Foth (2006) concludes:

> *If it is easy to set up mini case studies and initiate multiple micro action research projects within each case, the process of connecting the micro sites to a larger meta-network will contribute to exchanging valuable insights, experiences and narratives that ultimately promote action research as a viable research paradigm (Foth 2006 p222).*

Although the multiple case studies in my research are not "action research" (which would include iterative practice and research cycles), Foth's concept parallels my approach in aiming to design a standard methodology for team writing research, which can be applied to multiple case studies.

This research benefits from situated discourse and knowledge from the participants and provides in depth analysis of project activities in a way which can build a picture across projects to inform more broadly on professional practice. To incorporate the multiple realities stance for writing research concluded on in section 2.2.6, I use a particular type of discourse analysis, email content analysis.

Email communication is essential in, and integral to the process of networked team writing and provides electronic records of the communication, its transmission, receipt and reading. "The things that make up the social world - including our very identities - appear out of discourse", (Phillips and Hardy 2002 p2). Thus email records provide useful data with which to study the progress of documentation projects, without incurring additional effort for the subjects, and without researcher intervention. Data collection happens in real time during projects, transparently to the writers. The data are "true" in the sense that they represent real activities in context, and are recorded in real time. Permission to analyse records can be requested after a project has finished, eliminating any research- or researcher-influences on behaviour during the project. This transparent and accurate recording of communications has enabled particular kinds of writing research (Diaper 1993; Honeycutt 2001; Panteli 2004; Pendharkar and Young 2004; Te'eni et al. 2001; Vaes et al. 2002). Quantifying variables in communications, such as use of personal pronouns, which represent solidarity and pro-social behaviour, provides a basis with which to explore interpretations qualitatively with team members post-analysis. This simple data collection and analysis technique can be applied for the entire duration of projects and to entire project teams (Edwards et al. 2006 p175).

Roberts (2001) describes content analysis as "a mapping of non-numeric artefacts into a matrix of statistically manipulable symbols" (Roberts 2001 p2697), thus involving measurement rather than analysis. Some of the data from my email analysis is qualitative, and some quantitative. Meanings extracted from email content are units of words and interpreted to derive and code in categories, thus premised to a certain extent on "grounded theory" (Strauss 1987; Strauss and Corbin 1991), not in the development of theory, only in the interpretive coding into categories. Other data collection in the email analysis, such as the number of emails sent, size of emails etc. are objective, quantitative or positivistic. This is also true for data on age, gender, qualifications etc. elicited from the questionnaire. The research design thus incorporates triangulation, with different data sources from interviews, questionnaires and the email analysis, with both qualitative and quantitative elements.

3.4 Projects

3.4.1 Search for projects

To search for organizations and individuals who might be prepared to share their email records on writing projects, I published short articles describing the research in the news letter of the Institute of Scientific and Technical Communicators (Edwards 2005a) and in the on line journal of the Usability Professionals Association (Edwards 2005b).

I posted the same information on several mailing lists of the Society for Technical Communication:

- Belgian chapter;
- UK chapter;
- Usability Special Interest Group;
- Technical Editing Special Interest Group;
- Scientific Editing Special Interest Group.

I additionally posted the call for "participants" on the mailing list and website of the Teleworking Association (www.tca.org.uk) and to all presenters for the Language in the Media conference held in Leeds, 2005. I contacted fellow researchers, academics and professional acquaintances working on suitable types of projects. Consenting participants recommended third parties, who might permit my research, who I also contacted.

Many people kindly responded, but withdrew their interest at various stages. Four published authors promised to provide data and were later prevented from doing so by co-authors. Three organizations originally agreed, but later withdrew their consent. Organizations are reluctant to participate because it impacts time, resources and presents potential risks to the security of intellectual property and/or company confidential information.

Data from several projects were collected, but not all were analysed. Inclusion criteria were:

- Projects in which team members work across networks with fairly regular email communications;
- Projects with a team size greater than three and more than 100 email records;
- Projects completed within 18 months from the start of the data collection.

My search for contexts to research resulted in email records from three projects, one of which was from my own working environment:

- Software user documentation. I had collected my incoming and outgoing emails from a project for my MA research in 2001. The Managing Director gave me written permission to study my own working practices further for my PhD research.

- Hardware user documentation in a client-supplier relationship. I collected all emails on the project from the supplier and additionally a small number of email records which the client company had kept. The Managing Director of the supplier company gave written consent, having discussed the research and acquired consensus from both the client and supplier participants.
- Postgraduate course and handouts. I collected all project-related emails from the course coordinator. The Society Chairman gave his consent by email, and the Vice Chairman and supplier of the emails signed a written consent to analyse her incoming and outgoing emails on the project.

3.4.2 Keyware project background

Keyware was founded in 1996 and has since been divided into separate parts, some of which have been sold. At the time of the data collection, the company developed and marketed central authentication solutions for security in large enterprises, using state of the art technology in voice, face and fingerprint recognition. Central authentication implies using a single system to provide security for computer access, network access, building access and whatever else the company requires to remain secure across their entire organization. Thus, the security checks a company uses anywhere for any type of access, might use the single system then provided by Keyware.

The company had around 250 employees worldwide, having its main offices in Brussels, Boston, and Ieper. Keyware was growing rapidly through recruitment and acquisitions, and was introducing an entirely new field, biometrics, to the central authentication and security market. The founders of the company had direct contact with all their employees and involvement in all the company's activities.

The documentation process I studied was to create an administrator's guide for a particular authentication product designed for general use within any organization using computers. The manual was therefore intended for an audience with technical knowledge of software and hardware. The project ran from 4th January 2001 until 19th July 2001 and 20 drafts were completed prior to the document's release (Edwards 2001 p76). The team composition working on the project was new, the document was written in Microsoft Word and we used version management software to support product and document development.

My role was the technical author. (An overview of my own experience and qualifications is presented in appendix H.) I had no technical expertise in the product field, or in software development. My collaborative writing at Keyware was mostly asynchronous and sequential; we did writing tasks separately, passing the work from one collaborator to another. I worked with around a dozen colleagues, who were based in offices in Zaventem, Ieper, and Woburn, Massachusetts. Communications on the project were almost entirely by email, with three conference calls, four FtF review meetings and two introductory meetings with new team members. Emails were circulated locally between the Ieper and Brussels offices and also

between these offices and the US office (see Figure 3-1). Fourteen people communicated about the documentation in 295 emails, from which seven main actors were identified (defined by having written > 3% of the email communications). Experience, qualifications, role in the company and writing roles for the main seven team members are recorded in appendix I.

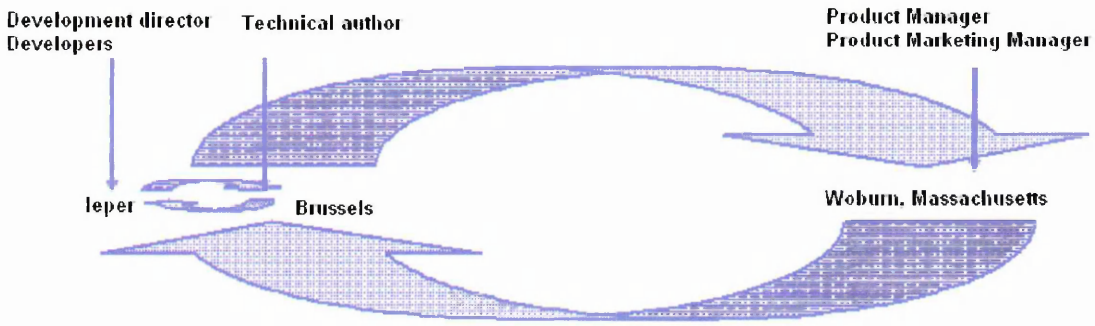

Figure 3-1: Keyware project: overview of communications between team members

3.4.3 Namahn project background

Namahn is a user-centred design consultancy based in Brussels, Belgium, supporting the development of media, products and tools. Examples of consultancy projects are requirements and interface specifications for software, improving the accuracy of on line form completion and designing and writing user documentation. Projects cross many sectors, such as finance, healthcare, IT and the public sector. Namahn's client for the project, Banksys, develops and maintains retail payment systems for Belgian banks, retailers and consumers. This involves management and development of payment tools such as electronic payment cards, and also support of the transaction system. Banksys payment terminals are used across Europe and the company has well over 1000 employees.

Banksys commissioned Namahn for consultancy supporting the design of the C-ZAM/XENTA™ terminal, which was launched in 2004. This terminal supports real time electronic payments by customers in shops. A bank card is placed in the device, which has a small digital display providing instructions. When prompted, customers enter their pin numbers to confirm payment for the displayed amount. Banksys also commissioned Namahn to write the C-ZAM/XENTA Service and Owner's Manuals. Operating internationally, the company commissioned the documents to be written in English. Namahn was specified to write these manuals conforming to the outline and template of a previous Banksys manual. The primary audience for the documents is international distributors, who can then adapt versions for local use (and language). The purpose of the Service manual is to guide distributors' engineers to install the units in shops. The purpose of the Owner's manual, which includes liability disclaimers, is to guide purchasers of the terminals, i.e. shop keepers to install the units themselves.

Most of the communication between team members on the project was by email, either because communications were between the client and supplier, who were based in separate offices, or because they were between the supplier and remote workers or subcontractors. There were three iterations, first, pre-final and final drafts, with one initial FtF meeting, and two FtF review meetings between the project manager at Namahn and the client team members. The illustrator attended two FtF meetings to review the drawings, one at Namahn and one at Banksys. Figure 3-2 shows a representation of the working infrastructure.

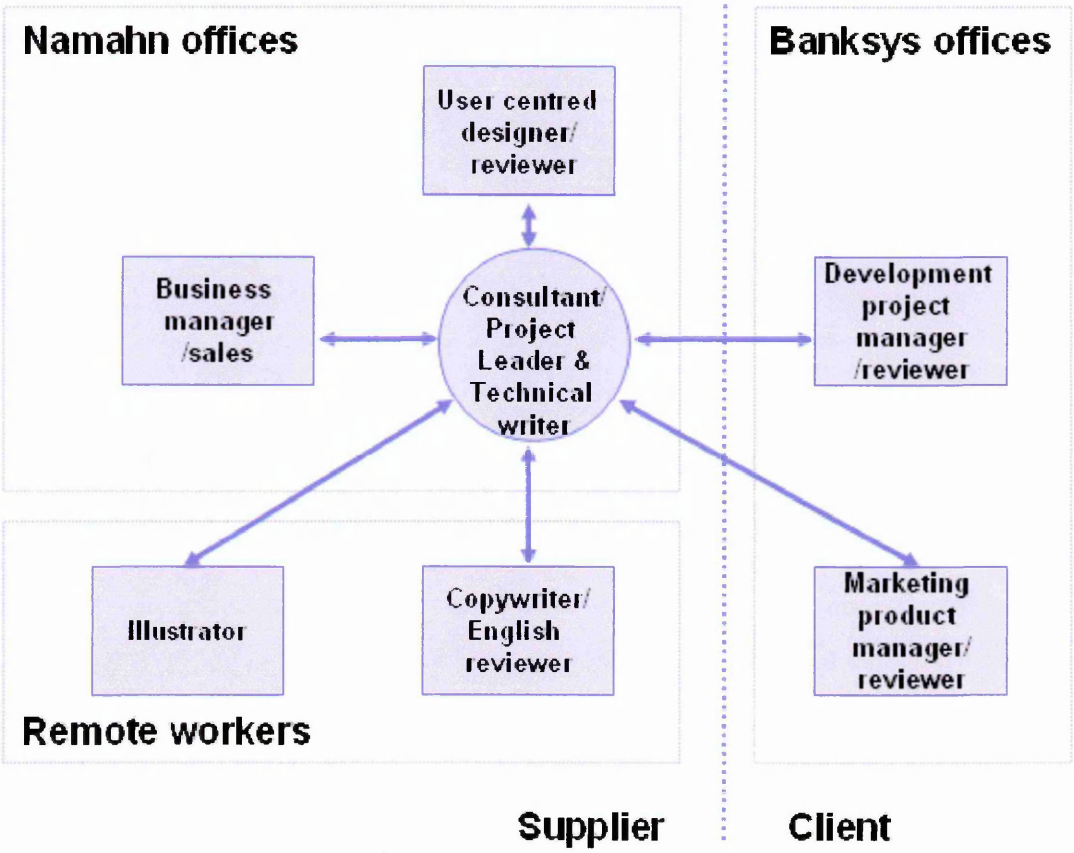

Figure 3-2: Namahn project: working and communication infrastructure

Team members on the supplier side (left in Figure 3-2) included the project leader, who was also the technical writer, a peer reviewer, illustrator, and a native English editor. Team members on the client side were the Product Manager (marketing) and the Project Manager (development). Team membership remained stable throughout the project, although the technical reviewer was the project leader during the first phase of the project, and had worked with the client team on software interface design prior to the documentation project. The technical writer then took over the leadership from phase two; he had not worked with the client team previously. The technical writer had worked with the reviewer at Namahn offices before, but not with the remote English reviewer or illustrator.

The project ran for two months from November 2003 to January 2004. The document was created in Microsoft Word. From fifteen people who communicated on the project, eight key actors were invited to complete the respondent questionnaire (described in section 3.8); respondent data is summarized in appendix J. I analysed 218 emails and defined the six main team members as authors or receivers of more than 10 emails.

3.4.4 Academic context: "the Society"

For the study of an academic context, email data was provided by the coordinator of a training course developed by a European Society which aims to promote knowledge, research and idea exchange on a clinical topic. The Society holds an annual scientific meeting, which is preceded by courses and seminars and followed by the Annual General Meeting. The Society is affiliated with the relevant scholarly journal. Approximately 1200 emails were saved by the main course coordinator, dated between February 2000 and May 2005, from 100 senders, and pertaining to the following Society activities:

- Course and seminar design, resulting in presentations and handouts for the course, which are being developed for publication;
- Review and selection of conference abstracts, which are published on line;
- Conference and course organization, involving the production of promotional material, the conference programme and development of a website;
- Annual General Meeting organization, resulting in agendas, minutes and newsletters for Society members;
- Other Society business.

These communications construct the discourse community in the Society and all pertain to document creation. I therefore consider them representative of the Society's communication style and have studied them all. The arguments for my research rationale demand this approach (see section 2.2.7 on including all the influencing factors in writing research, such as the organizational influences, history of interactions, and anticipation of future interactions).

Main team members were defined as individuals who wrote or received more than ten emails, which focused the analysis on 866 emails and 18 team members. Interviewee perceptions of team members' expertise and levels of socialisation are listed in appendix K (lines 45-46).

3.5 Initial interviews

Semi structured interviews with the project coordinators were used to gather background information on the Namahn and Society projects. Appendix L shows example questions and topics use for the Namahn interview, which was designed around the variables listed in appendices B, C and D, which are also summarized in appendix F.

In addition to responding to my questions, the project leader at Namahn also provided a PowerPoint file, which had been presented as a debriefing to the client at the end of the project. Further records which Namahn kindly provided included the emails and FileMaker records on the project and the documents themselves. I also gathered more information on the document audience, purpose and context of use in a phone interview with a key Banksys participant (see appendix L).

Project background information, course materials and post-course evaluations were kindly provided in the initial meeting with the coordinator of the Society project (see appendix M).

For the Keyware writing project, background information on the company, writing project and team members were sourced from the initial analysis (Edwards 2001).

3.6 Post analysis interviews

The research contacts at Namahn and Banksys were asked to respond to the preliminary results of the analyses in a telephone interview. Results from the email data analyses were listed in statements and simple graphical representations. These were sent by email to the interviewees to probe for their reactions. This member-checking technique aims to validate interpretations from the data with the feedback from the situational context, and thus improve the credibility of findings. The preliminary results and responses (in red) to these semi-structured interviews are presented in appendices N and O. These responses were validated by the interviewees as accurate. In the Namahn study, perceptions of greeting formality and social characteristics of the group were collected using a participant questionnaire, which is described in detail in section 3.8.

The same interview procedure was repeated with the Society Course Leader. I also asked the Society interviewee to rate team members on relative activity, involvement, sociability and formality. Further the Course Leader was asked to rate formality of coded greeting and signature categories (see appendix K lines 187-188) and social dimensions of the group (lines 111-112) using the same questions as used in the Namahn participant questionnaire (described in section 3.8). Interview data were transcribed and returned to the coordinator for validation, and are presented in appendix K.

3.7 Email data

3.7.1 Writing influences: choices and representations

Overview

In section 2.2, I showed how models of writing processes had evolved through different perspectives on research, textual, individual, group and social, and through different methodological stances from early exogenic to endogenic and interpretive stances. I concluded that the knowledge contributed by researchers towards an understanding of writing processes

has built on previous findings adding to and reinforcing earlier interpretations rather than challenging them. Further in section 2.2.7, I argued that professional team writing includes influences from all the perspectives, textual, individual, group and social, so that writing research should combine these to model the process realistically. Appendices B through E show the rich variety of variables to consider in writing research from the different perspectives. Appendix F maps variables into input, process and output categories, listing potential ways of measuring such variables. Appendix G categorizes in a slightly different way, considering how measures of these variables might describe the task and social dimensions of writing projects.

Table 3-1: Representing writing influences

Independent variable	**Representation**
Sender	*Keyware, Namahn and Society*: Individuals coded by role
Receiver	*Keyware, Namahn and Society*: Individuals coded by role
Direction Code	*Keyware*: Vertical organizational direction of email transmission: 0 = same level; 1 = up; 2 = down *Namahn*: Client-supplier direction of email transmission: 1 = client to client; 2 = client to supplier; 3 = supplier to supplier; 4 = supplier to client *Society*: Society and external organizational direction between roles of President (a single individual), Course Leader (a single individual), Academic Faculty members and external course and conference organizers. There were 14 possible directions (see appendix P), including two which are not presented in the results, in which an individual addressee is also the individual sender.
Distance	*Keyware*: Vertical organizational distance of email transmission. With 4 organizational levels, values range from 0 to 3.
Purpose	*Keyware, Namahn and Society*: Purpose of emails interpretively coded (see descriptions and intercoder instructions in appendices Q, R and S respectively)
Phase	*Keyware, Namahn and Society*: Cumulative email frequency category. Emails ordered chronologically and divided into six phases.
Dutch or English	*Namahn and Society*: Language coded: 0 = Dutch; 1 = English;

For empirical analysis of email data, I focused on a small number of variables, and collected other relevant variables through interviews and questionnaires (see sections 3.5, 3.6 and 3.8). In terms of a communication model for email communications, I extracted from each email the

writer, reader, purpose, and some indication of the writer-reader power relationship. Two further variables recorded were the socialisation phase, which allowed all other variables to be interpreted in the light of group evolution, and language, which controlled for the possibility that the mix of English and Dutch language emails might confuse interpretations. Table 3-1 shows the variables representing writing influences and gives a brief overview of how they are derived from the email data. The following sections explain in more detail the justifications for the representations and procedures for extraction.

Purpose

Earlier researchers have used several categorisations to differentiate purposes. In his analysis of emails from academic and commercial settings, Gains categorized main functions of emails into one of three types: informative, request or directive, after Ghadessy and Webster's classification (1988 cited in Gains 1999 p83). Sherblom in his early content analysis of emails received by a middle level manager in 1988 (Sherblom 1988 p46), coded the communication functions of emails as requests for information, providers of information, influence attempts, administrative items, or personal and social remarks. Dorn (1999) more recently surveyed 25 employees from a cross section of industries to explore the professional skills needed in business writing, resulting in purpose categories such as *to request action.* However, none of these categorisations are specific to writing teams; they allow coding of functions across disciplines, but are more abstract than the categorisation in my research.

As the objective is to study the email communication behaviour in the framework of, and to understand team writing practices, categories are interpretively coded according to contribution towards the team writing goal. This method was piloted in earlier research (Edwards 2001), in which I coded each email on a professional writing project in terms of activities which contributed to the main document goal of the project. Example purposes were *Review Discussion* and *Scheduling or Management* (please see appendix Q). This gave a clear overview of the types of activities completed during the project and, for example, the peaks in activities at certain times during the project. These profiles identified which tasks were completed by which individuals, and when, information which may help to predict additional resource requirements at certain phases during projects when plans impose critical interdependencies. Representing functional tasks in terms of relevance to the team writing goal thus adds value to the quality of information which may inform professional practice of team writing.

I coded purposes interpretively. For the body text of each email, I interpreted a main purpose of the email in terms of achieving the goal of the project. I created new codes as I found new purposes in the body texts. The coding structure developed for the writing influences is shown in Figure 3-3.

Senders, receivers and relationships

Adopting Nystrand's model of writing as a communicative event, and constructivist theory of communication, I chose to look at the power relationship between the reader and writer and how this influenced communication and social behaviour. Based on the "status equalization effects" afforded by email communications (Sproull and Kiesler 1986), such power relationships should not affect communication behaviour in emails; however, based on the writer's anticipation of the reader's needs and expectations (Nystrand 1989), such relationships should be reflected in communication styles. Organizational hierarchy has been shown to affect email style in earlier research (Sherblom 1988; Edwards et al. 2005 completed within the framework of this PhD program, also reported in chapter 4). In the Namahn (commercial) case study (reported in chapter 5), I analyse the client supplier relationship. In the Society (academic) case study, I analyse the organizational hierarchy, also with an element of the client-supplier relationship as represented between the Society members and external conference organizers.

Code System

Purposes
- Scheduling/management
- Review/revision discussion
- Circulation OR request for information OR content
- Accounts and or financing
- Document design
- Document/draft transfer
- Technology/application related queries/responses
- Politeness courtesy purpose
- Product related discussion
- Unknown purpose category/Other Society business

Individuals
- Senders
  - (by name)
- Receivers
  - (by name)

Figure 3-3: Coding structure for purposes, senders and receivers

I paid particular attention to identifying email recipients, as the "To:" field does not always accurately reflect the addressee. Studying "To:" and "CC:" fields in combination with the "Dear" field showed that very often the technical address of an email is convenience driven and does not match the personally named addressee within the body of the email. I therefore only used the "To:" field to identify the receiver if there was no personal open greeting. Multiple recipients were identified as such, rather than assuming that writers anticipated the first named receiver alone. References are made to the multiple sender characteristic of emails in the literature.

Baron writes "This blurring of distinctions between one-to-one and one-to-many dialogue was clear even from the inception of the technology [email]" (Baron 2001 p249). Nickerson also differentiated communication styles for single and multiple recipient styles in her research (Nickerson 2000 p156). I therefore additionally coded for one to one, one to one with an audience, and one to many in the Namahn and Society case studies.

Senders are more straightforward to code because technically there is only one "From:" identity possible. There was one sender in the Society case, however, who either dictated his emails to a third party, or used the third party's email address. As the original notions of communicative expressions were defined by the person dictating, I have treated these emails as though he had typed them personally.

Senders and receivers were manually coded in MAXqda for each email and exported to Excel. I assigned direction variables for each email record manually in Excel, based on the imported senders and receivers and information gathered on the organizational structures from the interviews.

Socialisation phase

Researchers have demonstrated the changing behaviours of teams both with and without computer support over time (Burke and Chidambaram 1999; Chidambaram 1996; Gersick 1988, 1989) and some criticism has been directed at research into computer supported teams, which has not taken a longitudinal view to allow for acclimatization to and reshaping of technology. Such research cannot therefore demonstrate social information processing or adaptive structuration theories (Chidambaram 1996, discussed in section 2.3.4). I do not consider the acclimatization of team members to email use as currently relevant, as most professionals using email in their work are expertly familiar with its use. However, changes in team behaviour and development of socio-emotional exchanges to help maintain the team are under focus in this research, for which reason the chronological order of emails needs to be taken into account.

In a virtual or semi-virtual scenario, on line communications shape the social structure of the team, rather than the daily FtF communications occurring in traditional working environments. Socialisation was therefore interpreted on a continuum scaled by the number of communications, divided into six phases for each project. I entered the date and time of each email manually into the attribute table of MAXqda to allow chronological ordering and phase coding.

3.7.2 Social behaviour: choices and representations

Overview

In section 2.3.2, I discussed the concept of task and social dimensions in team work and the value of the social dimension to team working. Further in section 2.3.5, I showed how strategies

are used in emails to communicate socio-emotional content. Socio-emotional communication behaviours in emails are used in this research to interpret team behaviour.

The dependent communication behaviour variables measured to represent social behaviour in this study are body text length, open and close greeting lengths, style of greetings, for which purpose a formality scale was developed, singular and plural first person pronouns, and social building units. Table 3-2 shows the variables representing social behaviour and gives a brief overview of how they are derived from the email data. The following sections explain in more detail the justifications for the representations and procedures for extraction.

Elaboration of body texts and greetings

Word counts for body text, open and close greetings represent writer's effort and value attributed to the communication in terms of the extent of their elaboration. Some researchers have studied length to compare email styles. For example, Crystal (2001 p115) has looked at variations in paragraph length between personal and institutional emails. Others have used word counts for their calculations of other variables such as % pronouns (e.g. Nickerson 2000). I've used word counts for both purposes, to compare communication behaviours and to calculate other variables.

Email greetings have been researched as interpersonal markers of style by several researchers (e.g. Crystal 2001; Danet 2001a; Gains 1999; Nickerson 2000; Sherblom 1988; Sproull and Kiesler 1986). Sproull and Kiesler write in their report of research on 1248 emails in organizational communication:

> *The total number of words in both salutations and closings is an indicator of the total attention paid by the sender to the social relationship. The number of words in the closing compared to the number in the salutation is an indicator of the relative focus on the self (Sproull and Kiesler 1986 p1500).*

In the Keyware study, I counted words manually to separate out texts written by previous and current email authors. Word count elaboration included greetings and automated signatures of up to 30 words. As automated signatures may not have been included intentionally (discussed in the following section), this artefact distorts the representation of value and effort attributed to the communication. Elaboration results from the Keyware study are therefore not analysed in depth in chapter 4, and I improved elaboration representation for the Namahn and Society studies. For these two studies, each email was coded in MAXqda as follows. First the text types were coded, such as open greeting, body text etc. (see Figure 3-4). This allowed automated word counts for each text type. Greetings were then coded in vivo, so that mostly the category name or subcode was the actual text, e.g. "hello". This type of coding leaves no room for interpretive error. Similar greetings were then grouped as appropriate during coding. Coded texts were then checked for consistency within each category.

Elaboration representation by greetings was also improved in these studies over the Keyware method. In the Namahn data, open and close greeting behaviour adopts certain within group norms (presented below in Figure 3-5 to Figure 3-6), and for most of the data (> 50%), greeting length only varies between 0 and 2 words. Representing greeting length as a percentage of body text therefore

distorts the representation. Percentages of body text will erroneously reduce the representation of a greeting in a longer body text, compared to the same absolute greeting length in a shorter body text. Absolute word counts for greetings are therefore used in the Namahn and Society case studies reported in chapters 5 and 6.

Code System

Text types
- Open greeting
  - (sub-categories = almost in vivo codes)
- Body Text
- Manual signature
  - Abbreviation
  - First name
  - Full name
- Close greeting
  - (sub-categories = almost in vivo codes)
- Automated signature

Social Markers
- Solidarity/cohesion markers
  - Dutch singular first person personal pronouns
  - Dutch plural first person personal pronouns
  - English singular first person personal pronouns
  - Dutch plural first person personal pronouns
- Social building markers
  - Apology
  - Courtesy
  - Humour
  - Self-disclosure
  - Social building units e.g. "Happy New Year"

Figure 3-4: Coding structure for text types and social dimension markers

Table 3-2: Representing the social dimension

Dependent variable	Representation
Word Count	*Keyware*: Total (semi-automated) word count for email. For emails with annotations of previous authors' content, new words only were counted manually. *Namahn and Society*: Absolute word count in body text of email. The editing function in MAXqda allows any content written by earlier authors and included in the current email to be separated out from newly authored content. By coding the body text part of each email, automated word counts were possible.
Greeting style/formality score	*Keyware*: Code assigned interpretively on a "continuum" of indicators from no greeting (=1), to conversational type greeting such as "Hi" (=2), conventional written greetings, such as "Dear/Best Wishes" (=3) to more formal greeting such as conventional written address and full email signature with organizational title etc. (= 4) *Namahn*: A score derived from the situated perceptions of the formality of four items: open greeting, close greeting, manual signature type and presence/absence of automated signature 0 = very formal; 4 = very informal *Society*: A score assigned interpretively to three items, open greeting, close greeting and manual signature. Each element was coded and the codes collapsed into four categories. These were then assigned scores for formality based on the informal spoken language and formal written language "continuum" concept (see section 2.3.6). Combined scores for the three elements ranged from 3 (very informal) to 9 (very formal)
Open greeting length	*Keyware*: Percentage of the total email word count used in start and end greetings combined *Namahn and Society*: Absolute word count of open greeting
Close greeting length	*Keyware*: Percentage of the total email word count used in start and end greetings combined *Namahn and Society*: Absolute word count of close greeting
% First person singular pronouns	*Keyware*: Frequency of first person singular personal pronouns and possessive adjectives as a percentage of the total email word count *Namahn and Society:* Frequency of first person singular personal pronouns and possessive adjectives as a percentage of the total body text word count
% First person plural pronouns	*Keyware*: Frequency of first person plural personal pronouns and possessive adjective as a percentage of the total email word count *Namahn and Society:* Frequency of first person plural personal pronouns and possessive adjective as a percentage of the total body text word count
Sociability score	*Namahn and Society:* Absolute frequencies per email of social building units including meaningful units representing apology, courtesy, personal disclosure or general social building units.

Based on interviewee perceptions of increasing formality with increased greeting length, I explored whether length of greeting might provide an indicator of formality (in addition to effort and value) in the Namahn data and present the argument for this here.

Figure 3-5 to Figure 3-8 illustrate the trends in open greetings, close greetings, manual and automated signature types. All three types of greeting show a definite "norm" in communication behaviour. In Figure 3-5, we see that the first three categories appear to represent the "norms" for open greeting. These emails account for 83% of the communications, while the remaining categories each account for less than 10%. Word lengths in these three categories, accounting for 83% of the emails, vary between 0, 1 and 2 words. This fact lends credence to the possibility that interpreting formality from greeting length in the Namahn case study is valid. No open greeting (word length 0) assumes a relationship which requires no greeting, in the same way that friends might start a conversation without a traditional open greeting on the telephone. A discourse context and shared understanding of the social dimension is assumed and no greeting is necessary.

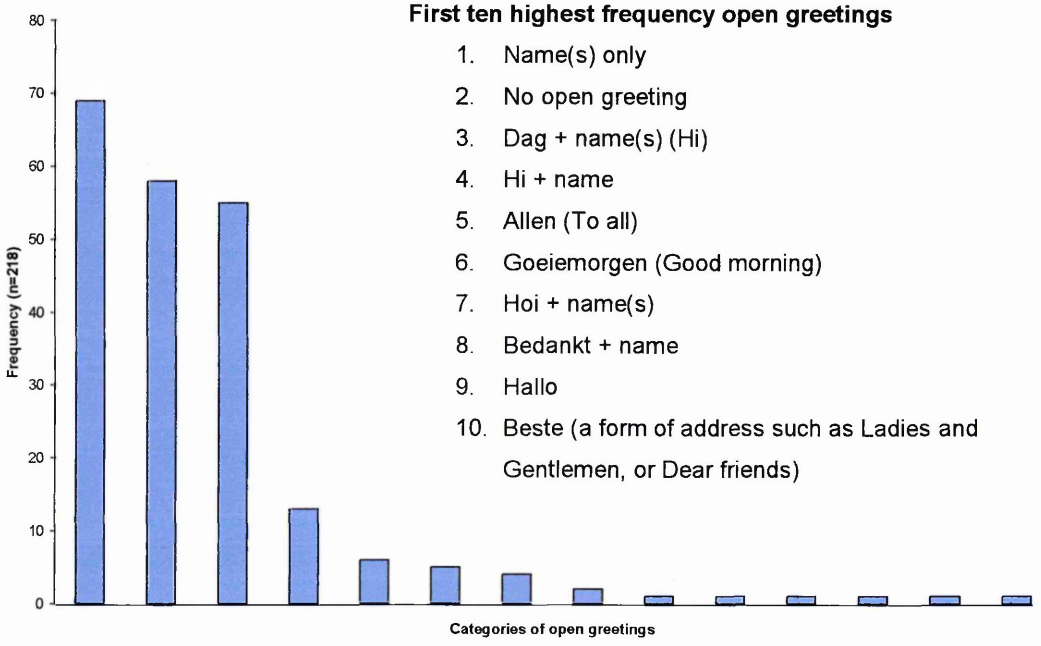

Figure 3-5: Namahn study: Open greeting frequencies (n=218)

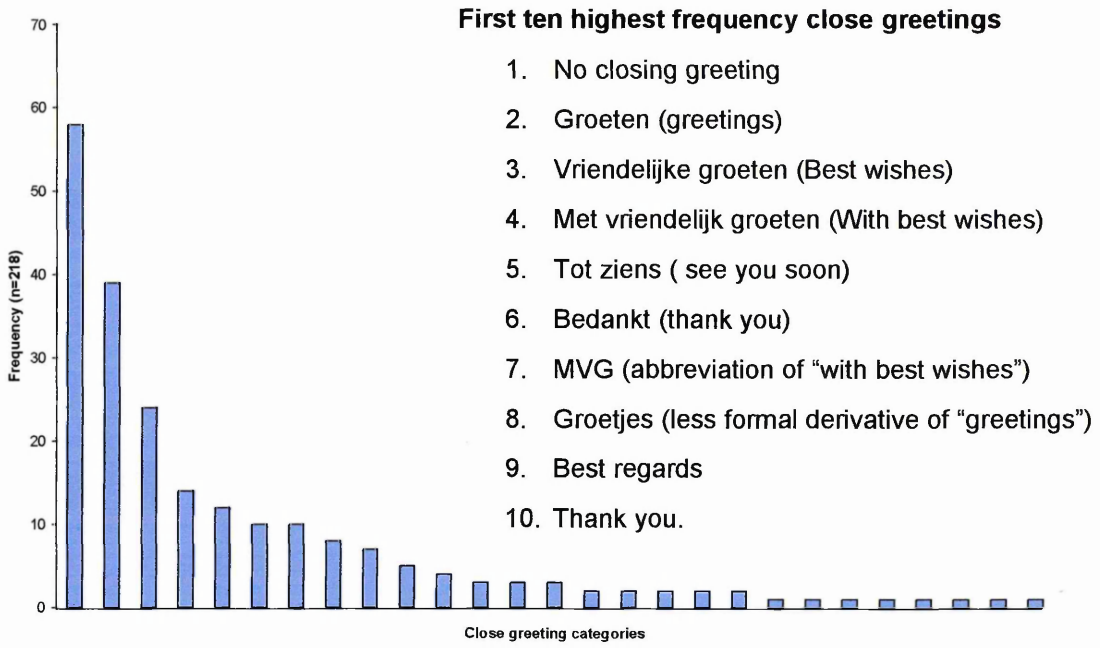

Figure 3-6: Namahn: close greeting frequencies (n=218)

Once again in Figure 3-6, we see a threshold for the most frequent norms of communication behaviour. The first three categories account for 56% of the communications and the remaining categories each account for less than 10%. The first three categories are no greeting (0 words) "Groeten" (1 word) and "Vriendelijk groeten" (2 words) (non-literal translations equating to "Regards" and "Kind Regards," i.e. the first is a less formal abbreviation of the second). Absence of a traditional close greeting address suggests an assumed relationship with the addressee, the abbreviated form shows courtesy, but is less formal than the full form of traditional close greeting. Thus the qualitative nature of the norms of close greeting behaviour demonstrates a scale of formality which mirrors the increasing word length of the greeting.

A similar pattern can be seen for the manual signatures in Figure 3-7. Out of the four types, only the abbreviations category accounts for less than 10% of the emails. The first three categories account for 94% of the communications and are full name (2 or more words), first name (1 word) or no name (0 words). Thus interpretations of the formality of the actual manual signatures used conform to the conclusion that increasing length represents increasing formality. Thus for the most frequent types of open and close greetings, and manual signatures, word length appears to reflect a scale of formality in the Namahn study.

An interesting result presented in section 3.9 is the similar positive correlation between the formality scores and close greeting in both the Namahn and Society case studies (see Figure 3-17f). Longer close greetings in both projects also occurred in emails which were rated as more formal. There was also a correlation between formality and open greeting in the Society data ($r=0.44$; $p= <0.0001$); emails with longer open greetings were rated overall as being more formal.

Thus, even though the scale of variation in greeting word length is very small, varying in the categories with the highest frequencies from 0 to 2, there is some meaning in these variations, i.e. such variations do represent variations in formality of communication behaviour. Interpretations from greeting lengths in the Namahn case study are therefore based on the assumption that longer greeting lengths represent higher formality.

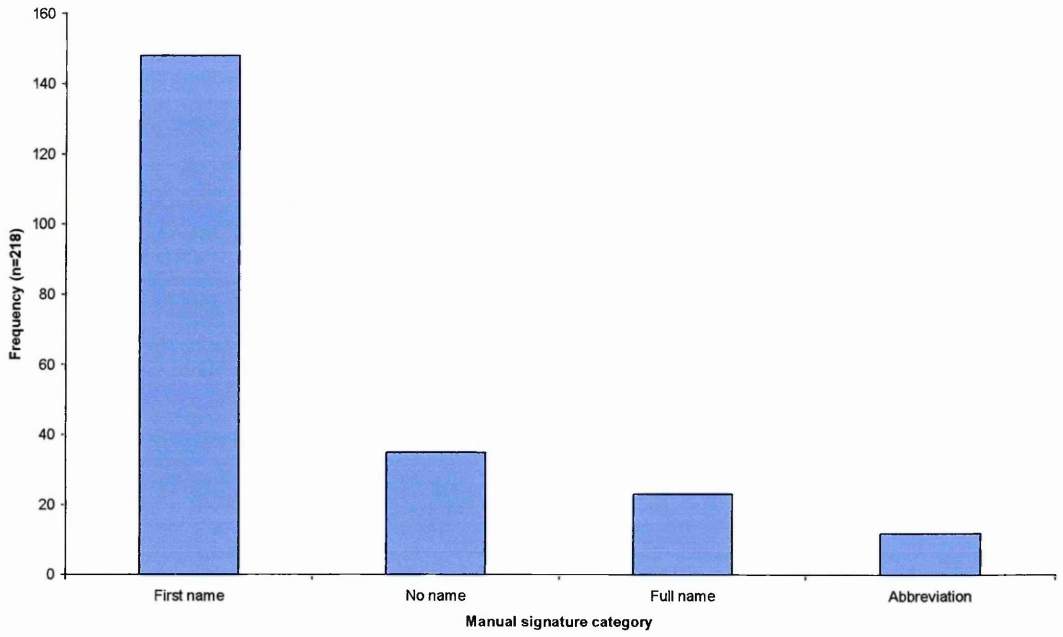

Figure 3-7: Namahn: manual signature frequencies (n=218)

Greeting style / formality score

Overview

In her research to explore the existence of, and tendency towards norms in public and business emails, Danet (2001a) coded open and close greetings and signatures by level of formality. By comparing emails to a business letter template, she found that "the greatest departures from the template were in the openings and closings" (Danet 2001a p77), which makes greeting style a particularly interesting and perhaps informative marker to study.

I explored two concepts to develop a formality score in this research, one based on situated perceptions (Namahn case study) and one based on a scale from less formal conversational style to more formal written style (Keyware and Society studies).

Formality based on situated perceptions

For the Namahn case study (chapter 5), formality of communication behaviour was calculated from scores of four items: open and close greetings, manual and automated signatures. Once I had coded the greetings into what were essentially in vivo type categories, I invited participants to judge the level of formality of the types of greeting and signatures used in the emails on the project. This provided a context sensitive evaluation of the level of formality, interpreted from

greetings and signatures, representing how senders and receivers in the team perceived the level of formality in their team's emails.

Occurrence of the manual signature types were scored with the average rating from the respondent perceptions, as were the different levels of greetings. Presence of an automated signature was scored as 3 (formal valence) and absence as 1 (informal valence). The scores available were averaged for each email and used as the formality score.

I treated the respondent perceptions (n=7) as multiple interrater codings. Kendall's coefficient of concordance (0= poor agreement; 1 = complete agreement) is recommended for evaluating the inter-rater agreement of rank data for more than two raters by Wuensch (2003). Kappa is used to test the null hypothesis that raters would agree by chance alone. Although the Kendall's coefficient of concordance was 52% ($p = 0.001$; appendix T), the data is essentially nominal, representing categories of formality, invalidating this statistic. The overall kappa value was only about 20%, or "slight" (Wuensch 2003), so that interpretations based on this formality score in the Namahn study require caution.

Low inter-rater agreement is possibly attributable to the inclusion of the automated signature as an element of the formality score. There is some evidence from Nickerson's research that automated signatures may not reflect the same style as the body text of an email (Nickerson 2000 p157). Inclusion of an automated signature may not be voluntary or intentional. Many organizations require a formal signature as an organizational norm, to present a certain image of the organization, its employees and professionalism. Also complicating the issue is the fact that inclusion or omission may not only lie outside the free choice of the author, but cannot reliably indicate intentional formality. If an unwanted automated signature is already included when an author starts writing a new email, the onus is on the author to remove it. Failure to remove an automated signature may not represent formal communication behaviour, but rather time pressures or forgetfulness.

In the Namahn study, I scored inclusion of an automated signature with a formal valence. Automated signatures with a manual signature were scored 1 or below (where 0= very formal and 4 = very informal) by six of the seven respondents and the remaining respondent scored "don't know". Thus presence of automated signature with a manual signature was perceived as formal. Automated signatures without a manual signature were scored 1 or below (i.e. formal valence) by five of the seven respondents and value 3 (informal valence) by two respondents. This may be due to the point raised above, that presence of the automated signature may simply be due failure to remove it (i.e. no action). Combined with no action in terms of including any kind of manual signature, two respondents felt this indicated informality rather than formality. Figure 3-8 shows that an automated signature was not present in the majority of emails (62%).

Further inspection of the raw data showed that the automated signatures were only present in supplier emails. This indicates a difference in communication behaviour norms between the client and supplier organizations. 48% of the total supplier to supplier emails had automated

signatures, compared to 61% of the total supplier to client emails, suggesting that relational direction may have played a role in the behaviour. For the supplier emails with automated signatures, 69% were "in house". However, "in house" on the supplier side includes two subcontractors, who worked remotely from Namahn (see Figure 3-2). Two out of the eight supplier authors who used automated signatures accounted for 83%, and these were the technical writer/project leader and the illustrator (subcontractor). Four receivers accounted for approximately 80% and these were the technical writer/project leader, illustrator, technical reviewer and client project manager.

Thus communication behaviour represented by automated signatures may be confused by organizational norms, individual norms and also by contextual issues such as relational direction. Nickerson concluded from her own research that the final close in the corporate emails was pre-programmed and varied in language use and formality. She writes: "code used in the text of the message had little effect on the code used in the [close], as might be expected in other forms of written business communication, such as a business letter" (Nickerson 2000 p157). These confusing issues need to be taken into account when interpreting the results of the formality score and suggest that automated signatures are not a reliable element for interpreting formality of email style. For these reasons, I only used the open and close greetings and manual signatures to derive a formality scale in the final case study on the Society project.

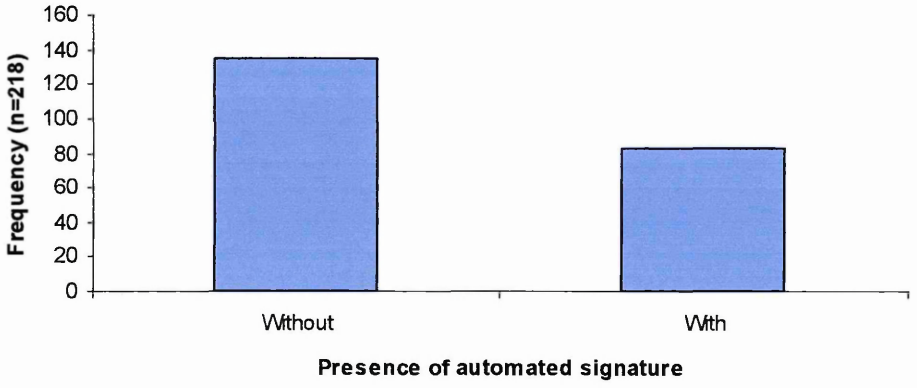

Figure 3-8: Namahn data: automated signature frequencies (n=218)

Formality based on a scale of conversational to written style

The second method of deriving a formality score, used in the Society study, is based on the concept of a continuum from spoken to written address. Many researchers are debating whether the language of CMC veers towards written or spoken discourse, or has perhaps developed a style of its own (e.g. Baron 2001; Ferrara et al. 1991; Harrison 2000a; Yates 1996; see also section 2.3.6). I assigned a formality score interpretively on a "continuum" from informal to formal. For open, close and manual signatures, I coded types into in vivo categories and collapsed them into a smaller number of categories grouping similar types, such as "Hi" and "Hey" for open greetings, and "Best Regards" and "Best Wishes" for close greetings. The

frequencies for each element of the formality score (open, close and manual signature), and the scores I assigned to each element are shown in Figure 3-9 to Figure 3-11. This method of scoring formality resulted in a scale from 3 to 9, with a reasonably normal distribution across the Society data, as shown in Figure 3-12.

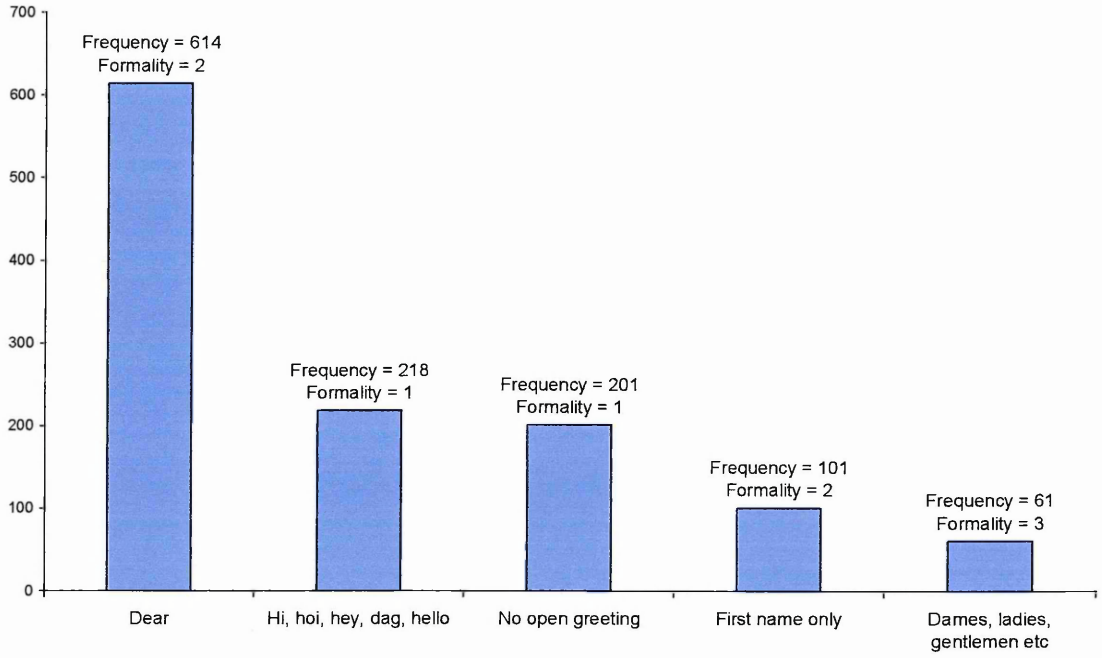

Figure 3-9: Society data – open greeting frequencies and assigned formality scores (n= 1195 emails; 1= informal; 3 = formal)

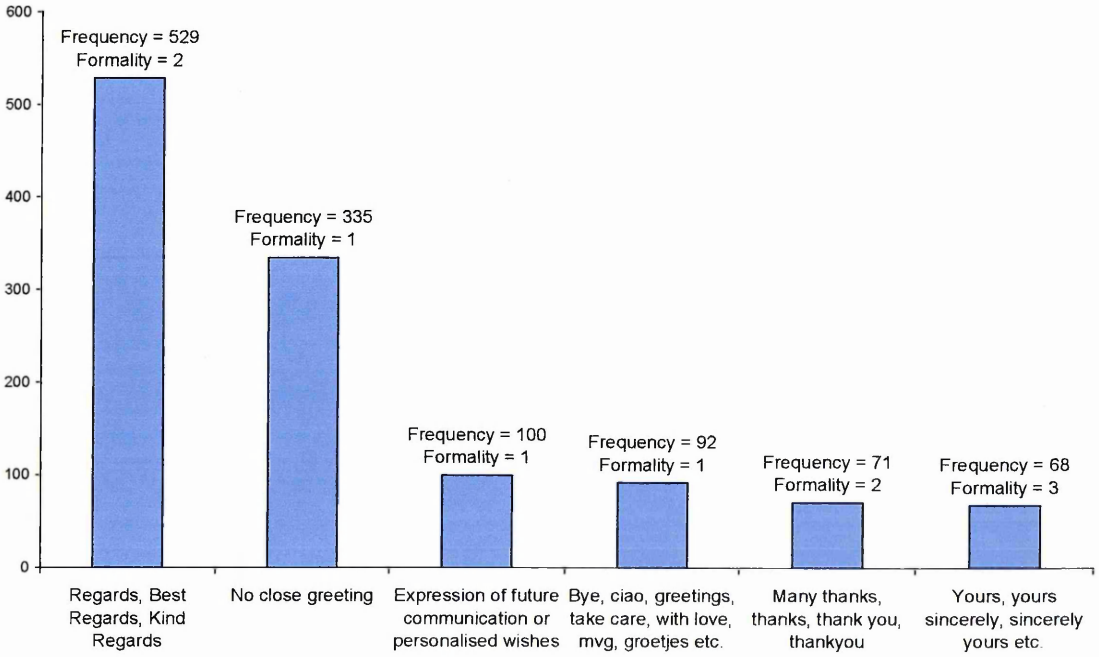

Figure 3-10: Society data – close greeting frequencies and assigned formality scores (n= 1195 emails; 1= informal; 3 = formal)

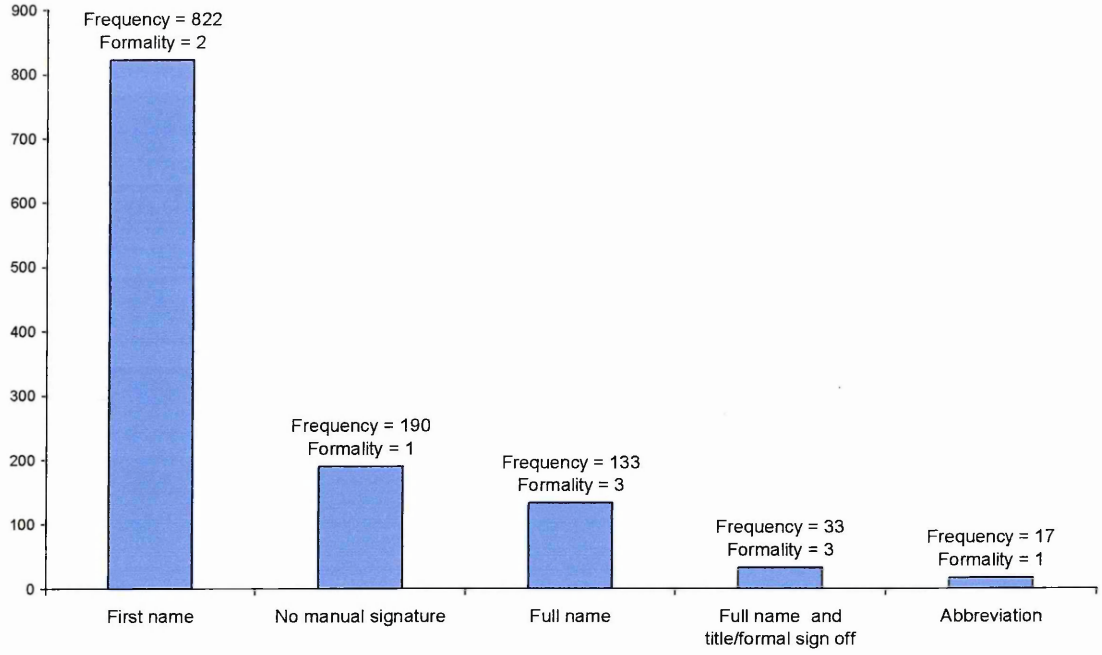

Figure 3-11: Society data – manual signature frequencies and assigned formality scores (n= 1195 emails; 1= informal; 3 = formal)

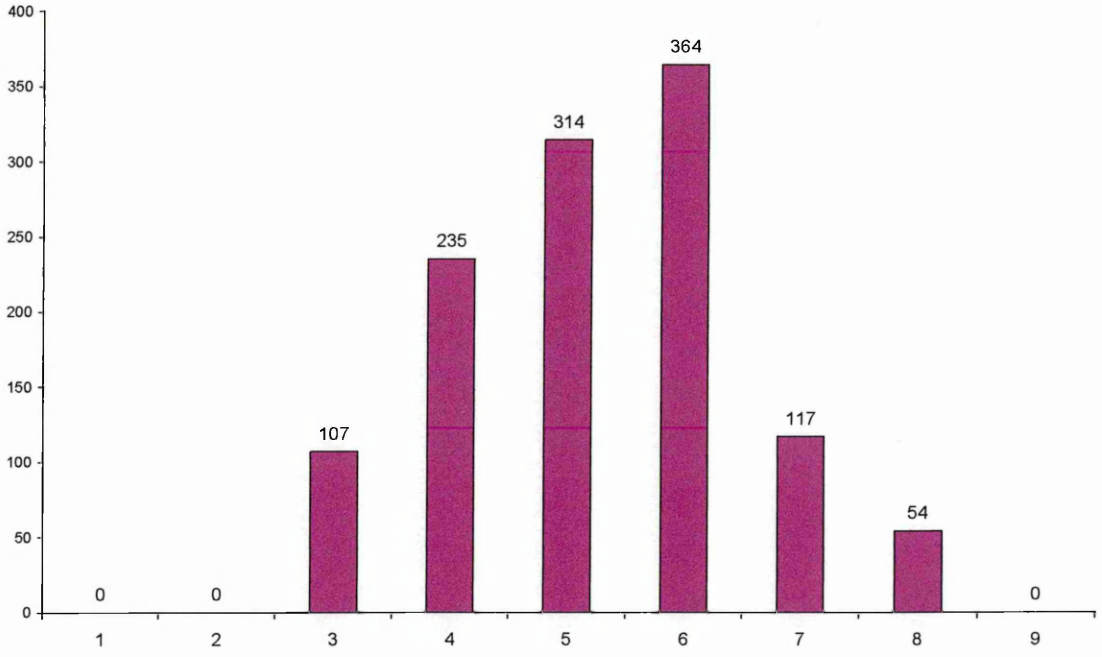

Figure 3-12: Formality score distribution in Society data (note that scores of 1 or 2 were not possible; the possible range is 3 to 9)

In the post analysis interview for the Society data (see appendix K), I asked the Course Leader to rank 13 team members from most to least formal. These team members were the most frequent email writers extracted from the preliminary analysis. She rated the President and AF5

as the most formal and AF2 and AF7 as least formal. Figure 3-13 presents formality scores derived from the email greetings and manual signature for these 13 senders, showing a conformity between the relative positions of these individuals on the formality scale derived from the emails and the perceived levels of formality by the interviewee, adding further credibility to the method of extracting formality scores from email content.

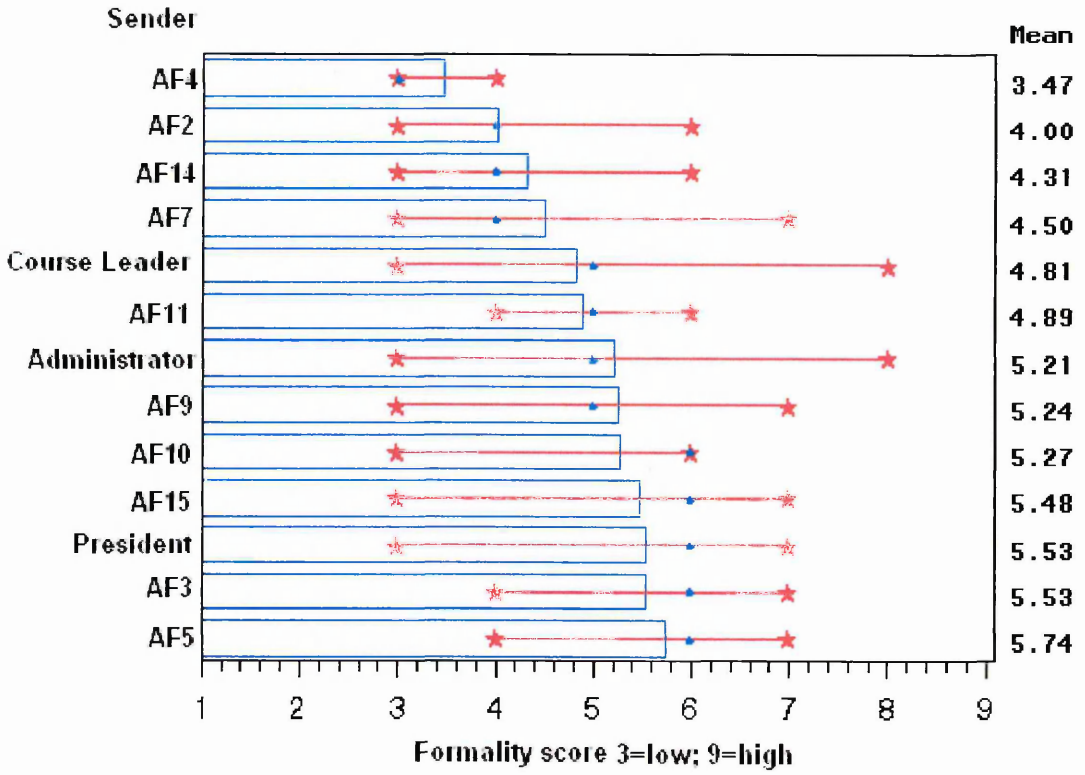

Figure 3-13: Increasing formality scores by sender for a subset of the Society data

Legend: min max median mean

First person pronouns

Nickerson writes:

> *...involved discourse highlights the relationship between the speaker/writer and their audience and the interaction between them, and the use of interpersonal rhetorical strategies contributes to the development of the relationship between the reader and the writer, and also indicates the writer's attitude to the propositional information contained in the text (Nickerson 2000 p168).*

These markers describe the involvement in the relationship between the writer and reader and the involvement with the topic, content or purpose of the email.

In a pilot study, Vaes and colleagues (2002) tested and validated the use of first person singular pronouns as indicators of different degrees of intimacy or solidarity in an interpersonal (email) context, where "mentioning oneself is to become more involved in the situation and to increasingly relate oneself to the other person" (Vaes et al. 2002 p527). Hyland (1998 p230) also uses first and second person pronouns as relational markers representing strategies to

involve readers and emphasize relationships in discourse analysis of organizational texts (although not emails); Eggins and Martin use "frequent references to the writer" as one of their characterizations of informal writing style (Eggins and Martin 1997 p232) and Postmes et al. used "self-references [I, me mine] as a measure of self-awareness" (Postmes et al. 2000 p349). Yates has shown higher first person pronoun use relative to total pronoun use in CMC than in writing or speech (Yates 1996 p42). With these interpretations in mind, and because counts of such units provide a practical (and possibly later automated) means of collecting ratio data, first person singular pronouns, possessive adjectives and reflexive pronouns *(I, me mine, myself, my)* were recorded to represent involvement, and first person plural pronouns, possessive adjectives and reflexive pronouns *(we, us, ours, our, ourselves)* were recorded to represent solidarity. Counts were used to calculate percentages of the total body text per email.

Correlation data from both the Namahn and Society case studies confirms the appropriateness of representing these markers as percentages of the body text rather than absolute counts (see appendices U and V). The positive correlations between body text word count in the emails and the first person singular markers for involvement (Namahn $r=0.36$; $p < 0.0001$; Society $r=0.72$; $p < 0.0001$) and plural markers for solidarity (Namahn $r=0.53$; $p < 0.0001$; Society $r=0.70$; $p < 0.0001$) justify using percentages of the body text word counts for these markers in the analyses rather than absolute counts. Otherwise counts would vary with the length of an email, thus distorting the representation of involvement and solidarity.

There is one common first person singular pronoun in Dutch and English ("me"), so it was necessary to differentiate between email texts of the two languages to count frequencies by language type accurately. Differentiation by language also allowed control for language dependent differences.

From a preliminary analysis of the Society data, I asked the Course Leader to rank 13 most frequent email writers from most to least involved. She rated AF3, AF4, AF15 and the President as least involved, and herself (Course Leader) and AF2 as the most involved (see appendix K, line 167-168). Figure 3-14 presents the involvement marker frequencies extracted from the email records for the 13 senders. Values represented by involvement markers conform to AF3, AF4 and AF15 being relatively less involved than others, and the Course Leader and AF2 as relatively more involved. The interviewee's perception of the President's involvement thus appears to conflict with the representation from the email data. However, there is some additional evidence to suggest that the interviewee's perception may be inaccurate. In the post analysis interview, the Course Leader was surprised to see a high frequency of emails originated by the President:

> *The President seems to have a large proportion. I think this may be because I was given responsibility to design the course and he had some expectations he wanted to communicate. That definitely reduced over time, so if you would look at the 2005 communications, it would be completely different (appendix K line 207).*

However, a plot of email frequency by sender with phase for the Society project does not substantiate the perception that the President's involvement in terms of email frequency decreased with time (please see appendix W, in which the President is represented by Sender 1). Indeed, the President's involvement both in terms of email frequency and relative first person marker frequency show relatively high involvement at a propositional and social level. This discrepancy between the interviewee's perception and the representations from the email data highlights a danger in interpreting involvement on one dimension, either with propositional content or with the relationship between the sender and receiver. The interviewee rated herself and AF2 as more involved than the President, and this perception appeared to be based on the task dimension, the goal of creating the course (see the quotation above). However, in email communications, first person markers represent involvement in the task, and also in the relationship in the interpersonal communication. The interviewee may have assessed her judgement on the task dimension alone, whereas the email data provides a picture of both dimensions. The President may show high relational involvement in his email communications, which is represented by the first person markers in the emails, but may not have been deeply involved personally in the course development, thus explaining the interviewee's perception. This highlights the care needed in interpreting first person markers in email communications with their dual representations of involvement in both the socio-emotional and task dimensions. This dual representation is discussed further in section 3.9 under "Correlations between social dimension markers".

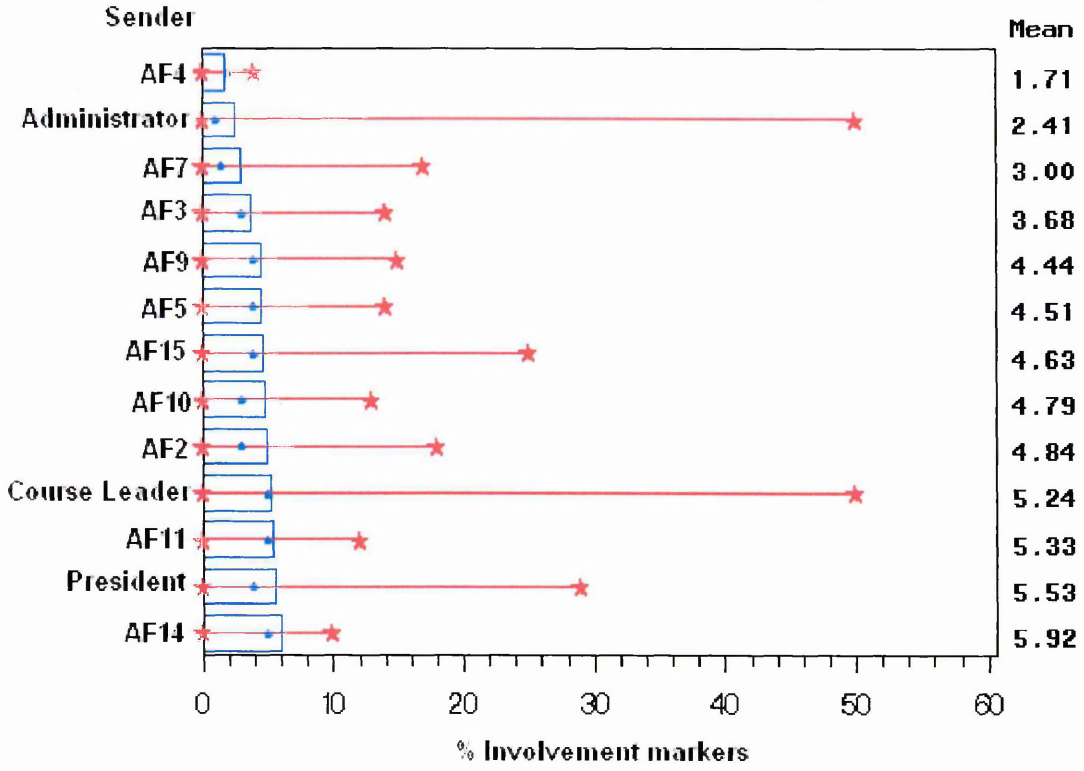

Figure 3-14: Increasing involvement scores by sender for a subset of the Society data

Legend: min max median mean

Social building units

> *Other areas of communication theorizing that might profit from incorporation of ...relational message themes include communication competence and relationship development (Burgoon and Hale 1984 p212).*

Social building units were interpretively coded from words or phrases which were non-task focused and which help towards building interpersonal relationships and therefore contribute positively towards the welfare of the team and its performance. The four main categories of social building units which emerged from the data were apology, courtesy, self-disclosure and general social building units. In particular, the degree of knowledge correspondents have of each other is an indicator of the intimacy or formality of the relationship (Argyle 1994 p132-134; Burgoon and Hale 1984 p204; Walther 1995 p191). Researchers argue that since self-disclosure is a verbal behaviour, it is less likely to be affected by the nature of the medium. For the coding I adopted Taylor's definition of self-disclosure: "any comments that revealed something private about the person" (Taylor 2000 p100).

For the Namahn case study, I coded these units interpretively as I went through the data. I coded meaningful units of text into categories for apology, humour, courtesy and general social building units, such as "have a nice week-end" or "Happy New Year". Once I identified a word as related to an apology or courtesy, I searched for the word and coded it throughout all the emails for consistency.

To improve on this technique for larger data sets, I devised a more objective and systematic protocol for coding socio-emotional content in the Society data. Using a list of all the words used in the body texts of the emails (a register of 9457 word types from a total of 116,164 tokens), I searched for words associated with the four social building unit categories. Reading through the list three times, I identified 280 word types or phrasal verbs from the corpus, which might highlight message content of apology, courtesy, self disclosure or general social building units. I used these words to guide my search and identify potential text units for coding, and coded the associated meaningful text unit into one of the four social building unit categories, as appropriate. The words used to search for potential social building units are listed in appendix X.

Words associated with apology included *sorry*, *apologies, apologise* etc. Words associated with courtesy were *please, thank you, grateful, appreciate* etc. Words associated with self disclosure were *children, disappointed, ill* etc. Searching for these words identified texts to consider. Texts which were not associated with the task goal, and which informed on something private to the individual were coded as self-disclosure. General social building units were searched for using words such as *congratulations, Christmas, excellent* etc. I coded text units as general social building, if the content aimed to strengthen relations by providing goodwill, support or sympathy.

Absolute counts of social building units were used rather than percentages of body texts, as the counts were low, ranging from 0 to 6 per email. Large variations in the body text lengths would therefore distort the representation of social building strategies, if percentages were used rather than absolute counts.

Again to explore the validity of using social markers in emails to represent sociability of individuals, I asked the Society Course Leader in the post analysis interview to rate the sociability of 13 most frequent email writers extracted from the preliminary analysis. The interviewee rated AF2 and AF9 and AF14 as the most sociable and AF3 and AF4 as the least sociable. Figure 3-15 presents increasing sociability by sender extracted from frequency of social building markers in the emails and shows similar rankings, with AF3 and AF4 having relatively lower means and AF2, AF9 and AF14 having relatively higher means.

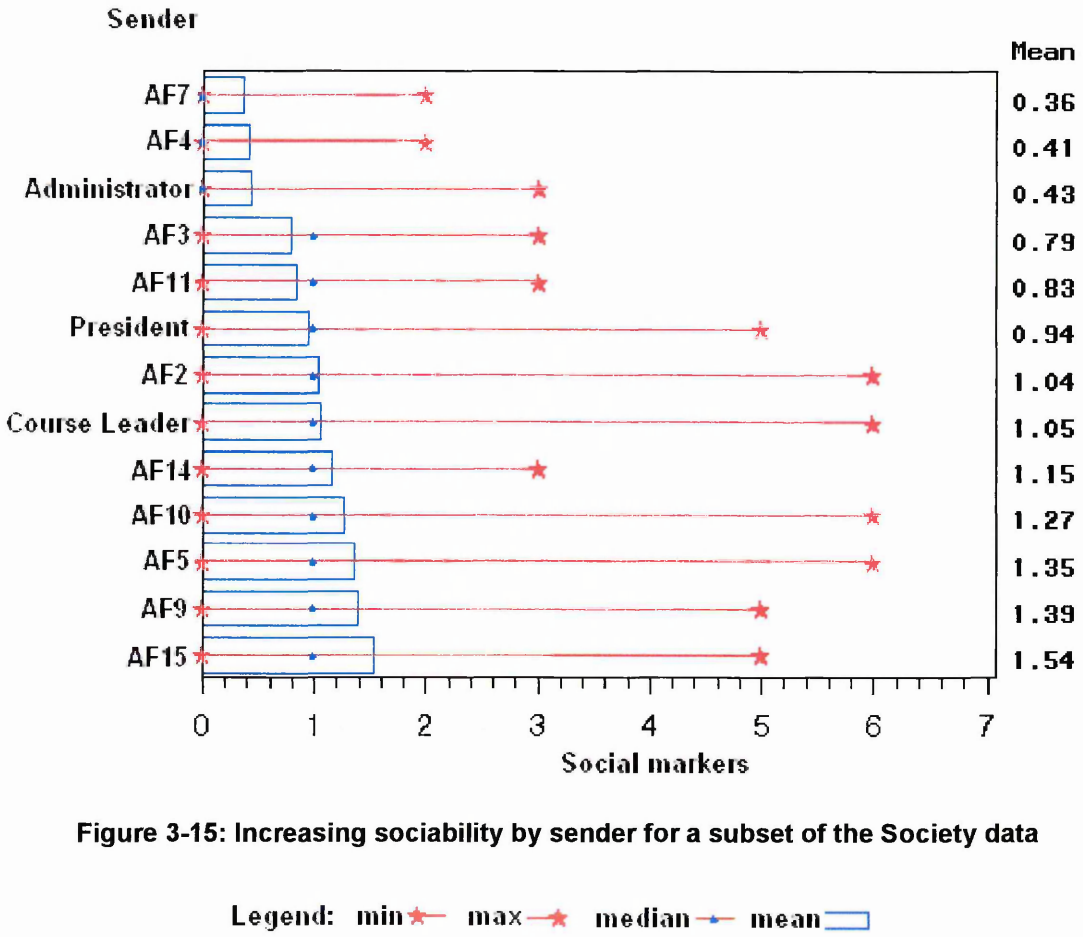

Figure 3-15: Increasing sociability by sender for a subset of the Society data

3.7.3 Email Analysis

Following coding of each email, email frequency and word count data for codes were transferred to Excel for calculations of percentages, and to develop further variables, such as direction codes. Category names were coded by number in the Excel file and consolidated into a single table for importing into SAS® (Statistical Analysis Software).

The dependent variables representing communication behaviour in the Namahn and Society case studies were analysed for correlations amongst themselves to further interpret the nature of their representations and the results are reported in section 3.9.

Although Analysis of Variance (ANOVA) is appropriate for multiple categorical independent variables, (Sharma 1996 p6), the computations for ANOVA rely on a number of assumptions.

The dependent variable should have a normal distribution. Homogeneity of variances is assumed, which means that the dependent variables exhibit equal levels of variance across the range of predictor variables; otherwise common within-group variances are inaccurately computed. Additionally the intercorrelations (covariances) should be homogeneous across the cells for the design.

Results from analysing the data with the SAS Univariate procedure (appendix Y) showed that the distributions of the dependent variables were skewed with high kurtosis (interpreted from values for skewness, kurtosis, box plots and probability plots). The "varmod.sas" macro (downloaded from support.sas.com) was used to test for homogeneity of variances. The SAS program and results of Levine's tests for homogeneity are shown in appendices Y, Z, AA and BB respectively. Unequal sample sizes may also affect the results. As this research is not experimental, but used data from real life working practices, the sample sizes per cell are unequal with many cells empty; this was easily explored using pivot tables in Excel.

I concluded that lack of homogeneity of variances, too many empty cells and the skewness of the data invalidated the assumptions of ANOVA. I therefore used non parametric tests, which do not depend on the above assumptions. The Kruskal-Wallis test is a non-parametric test used to compare three or more samples. It is used for one independent variable, e.g. purpose, with two or more levels and an ordinal dependent variable (UCLA 2006). It tests the null hypothesis that all populations have identical distribution functions against the alternative hypothesis that at least two of the samples differ only with respect to the location (median), if at all. These analyses are based on ranking of the data rather than the actual values of the data and are not dependent on normal distributions or equal sample sizes. If the p value is small, we can reject the idea that the difference in distributions is a coincidence, and conclude instead that the populations have different distributions. This is the analogue to the F-test used in analysis of variance, but does not depend on the samples having a normal distribution.

In SAS, the "Wilcoxon" option in the PROC NPAR1WAY statement requests an analysis using Wilcoxon scores. When there are two classification levels, this produces the Wilcoxon rank-sum test. For more than two classification levels, it produces the Kruskal-Wallis test (appendix CC).

The Friedman test would have been appropriate for an experimental design where individuals each provided a single value for each condition or class level, such as purpose category. However in this analysis, each email provides values for dependent variables for each class level. Each email is considered "unrelated" (Greene and D'Oliveira 2006 p73) because each email entity is not subjected to identical comparable conditions. If I had designed an experiment in which I asked individuals to write three emails, each with one of three purposes, the design would be related, with dependent variable values for each of three conditions for each subject. This related design would need to be analysed with control for the subject, who was the same for each of the treatment conditions. In this research, the data being analysed is a sample of emails. Each email provided a set of values for the dependent variables. The same email was

not sent for different purposes, or to different recipients. Thus each email is considered unrelated.

SAS also offers an option to compute the exact p values for the Kruskal-Wallis test. This computation is recommended to avoid drawing conclusions from calculations of approximate p values which may vary, possibly only very slightly, but still affecting significance thresholds (Narayanan and Watts n.d.). Exact computations of p values, however, require high computational resources and were not possible on my own computer. K.U.Leuven (where my local supervisor, Professor Arthur Spaepen is based) kindly offered access to their LINUX Cluster, for which I wrote a shell job script to send the SAS task and data for computation of exact p values. Various job scripts were tested with memory and time resources up to 10MB and 48 hours permitted, but did not compute the exact p values. An example job script for this purpose is shown in appendix DD. The data sample was reduced to 20, 30 and 50 lines (email records), which also returned the analyses without the exact p values. There is no way to estimate the time required for computation of the algorithm used for exact p values, and the resources required can be prohibitive (SAS Institute Inc. n.d. NPAR1WAY Procedure). I therefore concluded that the computational resources required for this exercise were too high for the scope of this research.

Finally for the Namahn case study, I ran the dependent variables for each email through factor analysis using SAS to explore for social constructs in the data. However, the correlations between the variables were low and the sample adequacy measured by Kaiser-Meyer-Olkin's measure (KMO: Sharma 1996 p116) was too low suggesting that factor analysis was inappropriate. I repeated this procedure in the Society case study, but this data also produced a low KMO measure (0.67= mediocre, Sharma 1996 p116).

Significant results from the Kruskal-Wallis tests are reported for each case study in chapters 4 to 6. Results of the Namahn and Society studies in chapters 5 and 6 are interpreted in the light of qualitative data collected from the interviews (sections 3.5 and 3.6), and additionally participant questionnaire data for the Namahn study.

3.8 Participant questionnaire

One of my aims with this research is to design a data collection method which can be used consistently across different contexts without disrupting work or introducing research or researcher influence, but which is context-sensitive, informed by the situational context for interpretations. For the Namahn data set I had access to the team participants and to avoid unnecessary disruption of working practices, I collected data by means of a short questionnaire (appendix EE). I sent this (by email) to the eight key email communicators on the project to complete when convenient. In this section, I describe the aims and design of the questionnaire.

There were three main purposes to the questionnaire:

- To collect individual data such as professional experience, project role etc. to qualify the relevance of responses. Section 1 of the questionnaire addressed this purpose.
- To gather perception data which could be used to fine-tune, and/or test interpretations from analysis of email data. Sections 2, 4 and 5 addressed this purpose.
- To collect participants' perceptions on the success of the project (a part of the research framework which is reported in chapter 7). Section 3 addressed this purpose.

The questionnaire was designed based on recommendations from the literature (Anderson 1985b; Denscombe 1998; Dillon 1990) to help the respondents complete it as quickly as possible, with checkboxes, pull down menus and text entry boxes in a form. Only a brief explanation of the purpose of the questionnaire was given as all participants were already aware of the research (reported in section 3.4.1). Instructions for completion and return were given, together with confirmation of personal confidentiality.

Section 1 of the questionnaire (see appendix EE) validated the individual's eligibility to respond to the questions and also provided useful data on their role in the project. This data is summarized in appendix J. Section 2 collected respondents' evaluations of the relative importance of different purposes that I had interpreted from the email data. This information provides a weighting to fine-tune the representation of project activity by email frequency (see section 3.9). Section 3 of the questionnaire gathered participants' perceptions on the project. Questions addressed project, personal and organization goals. The latter monitored organizational elements such as budgets, deadlines and performance relative to other projects in the organization. These data were further qualified by probing the respondents' judgement of the value of these aspects and are reported in chapter 7.

Section 4 of the questionnaire probes the social dimension of the project, such as group belonging and cohesiveness. The first five questions were adapted from Seashore's group cohesiveness index (Miller 1991 p375). The validity of perception data is particularly difficult to evaluate. For example the relationships between individuals or between organizations may inhibit individuals from reporting any lack of group cohesion (Beck 1993 p109; Kelly and Duran 1985 p191). Triangulation with email data in this research thus seeks to reinforce the findings of qualitative data by quantitative data and vice versa.

Section 5 listed all the types of open and close greetings, and signature types (manual and automated) and asked for respondents' opinions of their level of formality on a five point scale, 0 and 1 with negative valence, 2 as don't know or neutral, 3 and 4 as positive valence). This provided a score for each of the open and close greetings, manual and automated signatures, with which to evaluate the formality of style in each email (see section 3.7.2, "Greeting Style / formality score", "Formality based on situated perceptions").

3.9 Reliability and validity

Email coding reliability

MAXqda has several cross-validation checks for integrity of coding and results. One method is the conversion of codes to variables. The converted variable records the presence or absence of the code for each text, making it possible to check that all texts have been coded. Whenever I transferred variable data from MAXqda, to EXCEL, I also copied the text reference number and cross checked with the text reference column in the Excel file for consistency of data ordering. Retrieving coded segments in MAXqda allowed consistency checking within coded categories. For example, by activating and retrieving all the segments coded for *Apology*, I could run through these texts alone and confirm the consistency in my interpretation of *Apology* text units. All codes were checked in this way.

To test the reliability of the interpreted codings for *Purpose*, a random sample of 10% of the body texts were extracted (i.e. the body texts only, no names of individuals) for coding by a second rater, who first signed a confidentiality agreement. I used definitions of my categories of purpose as instructions for the second coder and these are in appendix R. A contingency table was drawn up to record the frequency of agreements between the two raters and the Cohen Kappa statistic analysed using Excel and Analyse-it applications. Inter-rater agreement was high with a Cohen Kappa score of 0.77 ($p < 0.001$), which equates to "substantial agreement" (Landis and Koch 1977 cited in SAS Institute Inc. n.d. Sample 507). Kendall's coefficient of concordance was 0.93 ($p < 0.001$: see appendix FF).

Because I manipulated data to collect and create variables between MAXqda, Excel, SAS and Word, I completed final integrity checks on the data submitted for statistical analyses. For 5% of the Keyware and Namahn emails and 2% of the Society emails, I checked all of the variables, reversing their codes to identify and check their values in the original emails. No errors were found in the Namahn data; error rates of 3% and 1% were found in the Keyware and Society data respectively. Errors found were miscodings rather than data ordering errors, i.e. all other variables for an email were correct apart from the error found. This served as a horizontal check (per email) for data integrity. Vertical checks were also completed per variable, as described above.

Email representation of activity validity

In this research I interpret activities through the lens of the email communications on networked team writing projects. The research assumes therefore that 1) email communication frequency represents activity, and 2) that the emails available represent a sufficiently large enough part of the total project communications to accurately represent activities.

On the Namahn and Society projects studied in this research, interview respondents estimated that 50% and 80% of the communications were mediated by email respectively (appendix N,

line 114; appendix O, line 99; appendix K, line 60). In particular the Namahn interviewees commented as follows:

> *Usually everything was followed up by an email and every meeting was prepared by an email. I would say 50% [communications by email] (supplier: appendix N, line 113).*
>
> *Any tracking goes on by email though, particularly with a documentation project, where the document is changing. I would estimate that at least 50% of the communications were by email (client: appendix O, line 97).*

These responses tend to validate the interpretation of email frequencies representing activity, particularly on documentation projects. Additionally, 50% and 80% are large percentages of the overall project communications.

Regarding the second point, i.e. sample adequacy, emails were provided by the main coordinators for all three case studies. As the main coordinator of a project is the most involved in project communications, these individuals' records are the most comprehensive and therefore most representative of overall project communications.

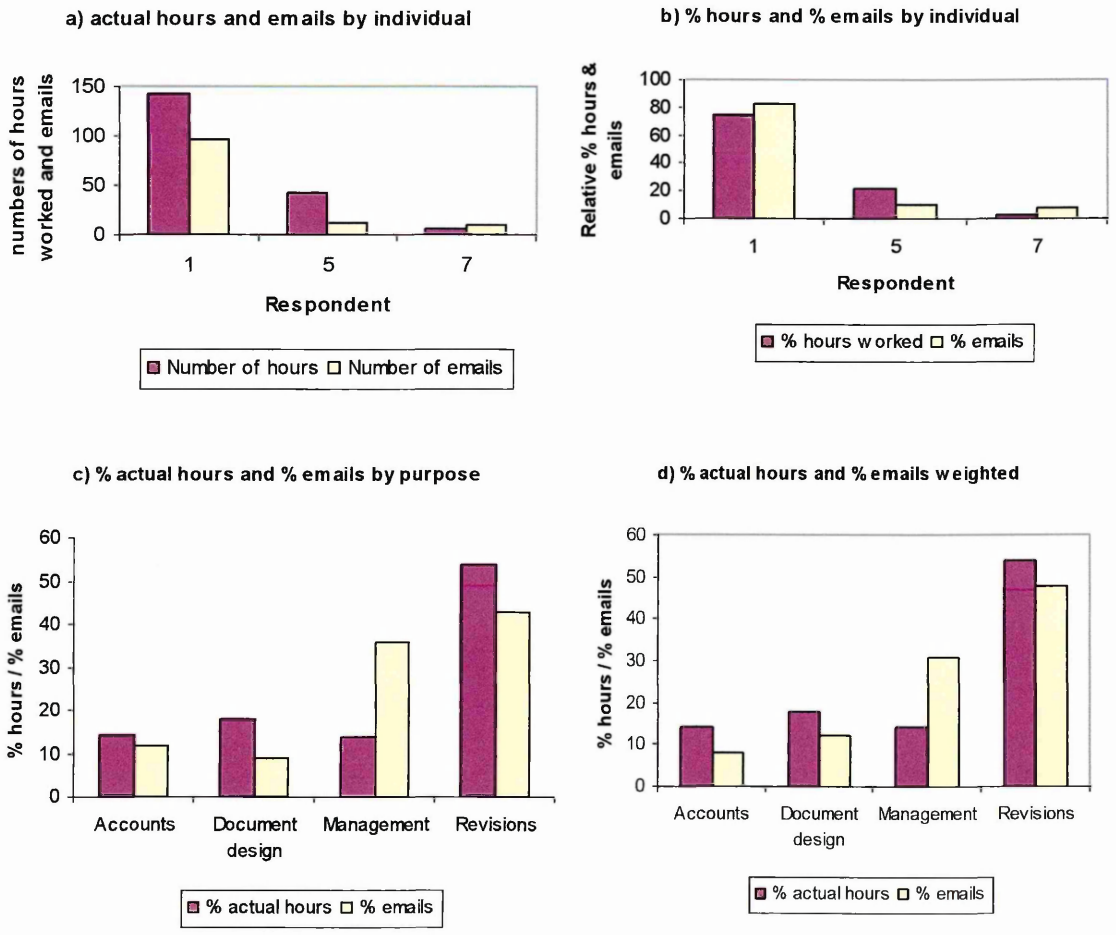

Figure 3-16: Namahn project: Comparisons between actual hours worked and email frequencies by respondent (a & b) and by purpose (c & d)

Comparing relative hours worked against relative email frequencies using records of consultancy hours from Namahn also validates the assumption of relative communications representing relative activity. Figure 3-16a&b shows that relative email frequencies for three Namahn participants show a similar profile to the relative hours worked by the participants. Profiles of hours and emails by category of purpose (Figure 3-16 c) are less well matched. Team participants (n=7) were asked to rate on a 5-point scale (very low, low, don't know, high and very high) their opinion of the value of each purpose towards achieving the goals of the writing project. Three respondents gave the same rating for all purposes and were excluded from the weighting calculations. The remaining four respondents' values were averaged and used to weight purpose frequencies of emails to test whether this improved email frequency representation of activity. Weighting with respondent perceptions of value of different purpose types corrects the differences very slightly (Figure 3-16 d), although inter-rater agreement between the respondents scorings on value were inadequate (Kendall's coefficient of concordance 0.37 $p = 0.148$). Mappings by purpose show the largest discrepancy in magnitude for the *Management* category, and different rankings in magnitude, apart from *Revisions*, which ranks highest for both hours and emails. Although time records were only available for three participants, and this study analyses only one small project, the similar mappings between hours and email frequencies reinforces the interviewee interpretations of email frequencies being representative of project activity.

Correlations between social dimension markers

Overview

As a preliminary exploration of overall differences in patterns of behaviour in the Namahn and Society case studies, and of the communication markers and their representation of pro-social behaviour, I ran a correlation analysis to search for differences and similarities in the two data sets. The complete results from the SAS computation of correlation coefficients for the Namahn and Society case studies are presented in appendix U and V respectively and are represented visually in Figure 3-17.

Correlations between social behaviour markers

First of all Figure 3-17a) shows correlations between the social behaviour markers, which vary for the two projects. Involvement and sociability have a small positive correlation for the Namahn project and larger correlation for the Society project. The involvement marker represents involvement in the task and the socio-emotional content of the message, so this difference could reflect that only the task component is represented in the Namahn data and I return to this point in section 7.3.2.

Small positive correlations exist in the Society project between all the remaining social behaviour markers, whereas for the Namahn project the only other significant correlation is between formality and sociability. In the Namahn data, sociability decreased with formality, whereas in the Society data, sociability increased with formality.

The difference between the two projects may be explained by the dual representation of social building units. These units represent pro-social strategies to build relationships, and are therefore expected to occur in communications between less familiar individuals, who therefore adopt more formal styles. They also represent the status of existing relationships between more familiar individuals, who therefore write less formally. Thus, individuals in the Society project used social building strategies to develop relationships, whereas in the Namahn data, social building units represented the closer less formal relationships.

This dual representation may be explained by the varying components of the social building units coded in this research and their potential dual representations by component. Self disclosure about family or other personal circumstances may be a norm within an established relationship, or a strategy to establish a closer relationship (Argyle 1994 p162). Courtesy strategies may be more common in less familiar more formal relationships. For example, "Thank you in advance for your time" might be a formal strategy used in professional communications. The frequencies of different types of social building unit, such as apology or courtesy, within the total count for social building units may therefore affect correlation of this marker with other socio-emotional communication markers, such as a formality. Dual representation and the different types of social building markers require further research, which I discuss in chapter 8. These issues also support the need to use several different types of markers, (such as first person pronouns and greeting styles) to interpret email styles and communication behaviours accurately.

Word count markers: body text, open and close greetings

There were also differences between the two projects regarding correlations in the word count markers (see Figure 3-17b). There were significant positive correlations between open and close greetings on both projects. Thus emails with longer open greetings also tended to have longer close greetings. Body text and close greeting length also correlated positively, although this correlation was small in the Namahn data. Body text correlated positively with open greeting in the Society data, but there was no significant relationship in the Namahn data.

Correlations between social behaviour markers and word count markers

Figure 3-17 (c) and (d) show that body text length correlated positively with involvement and solidarity in both case studies. Thus longer emails showed higher involvement and solidarity. There were no significant correlations between body text length and sociability or formality in the Namahn project, whereas in the Society project, longer emails showed higher sociability and also higher formality.

Figure 3-17 (d), (e) and (f) show that in the Namahn data, open greeting behaviour did not correlate significantly with solidarity, sociability or formality, suggesting that rather than being influenced by the social relationship between writer and receiver, other factors are more influential on the norms in open greetings in this project. The positive correlation between the

involvement marker and open greeting may be explained, therefore by the task component of this marker as opposed to the socio-emotional component.

In the Society data set, open greeting correlated with all the social markers, as shown in Figure 3-17 (c), (d), (e) and (f). Thus in the Society context, open greeting behaviour varied with the relationship between the writer and receiver.

Close greeting, on the other hand, correlated positively with all the social behaviour markers as shown in Figure 3-17, (c), (d), (e) and (f) for both contexts. Thus in both projects, close greeting behaviour varied with the relationship between the writer and receiver.

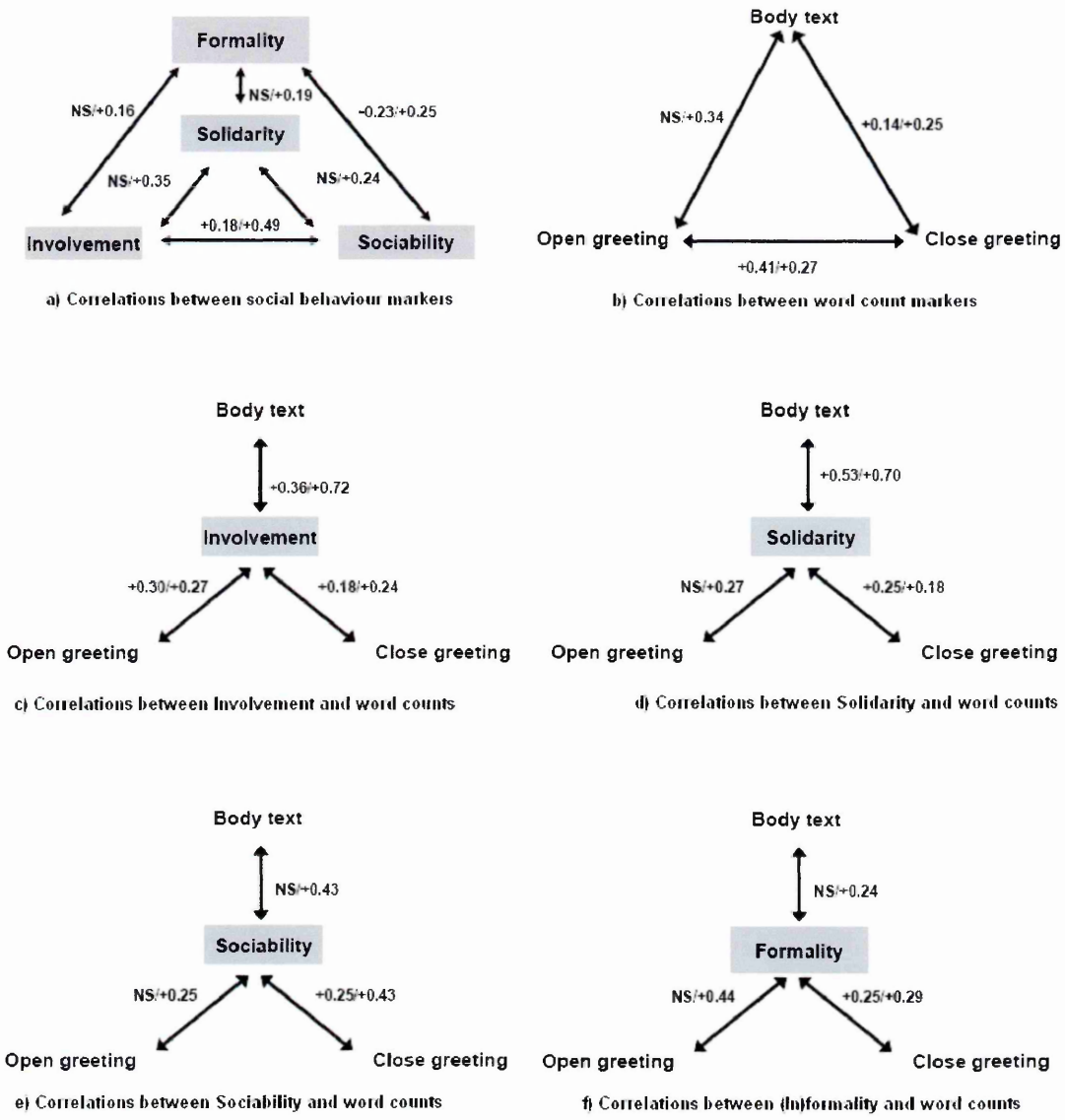

Figure 3-17: Significant (α=0.05) Pearson Correlations Coefficients between dependent variables (Namahn data/Society data)

Conclusions on correlation analysis

I completed this correlation analysis to review differences between the Namahn and Society case studies and to better understand the communication markers and their representation of social behaviour, so there are conclusions to be drawn on both issues.

Firstly there is some evidence to suggest that individuals in the Society project used social building strategies to develop relationships, whereas in the Namahn data, such representations were only used once relationships were established. Attention to the receiver given in open greetings in the Namahn project varies with factors other than socio-emotional relationship between the writer and receiver, whereas in the Society data open greeting correlated positively with all the social behaviour markers, suggesting that open greeting style varies with the relationship. Attention given in the close greeting correlated with all the social behaviour markers, suggesting the influence of the writer receiver relationship on style of this component of the email in both projects. Differences between the two projects are discussed further in chapter 7.

Each of the markers represents slightly different and sometimes multiple aspects of socio-emotional communication. The markers may also represent both task and social aspects. Interpretation therefore requires consideration of all the markers and combining the evidence they provide.

Strengths and limitations

In section 2.2.3, I criticized Flower and Hayes' work for the researchers not acknowledging that they were part of the social reality they were analysing. A particular strength of collecting recorded emails lies in its exogenic research perspective; subjects completed their projects, and were only requested retrospectively to provide their email communications for analysis, so there was no shared social reality between the researcher and the researched during the project, and therefore no researcher- or research- influence on the behaviour of the researched.

However, there is a risk of decontextualised interpretation of written content. In formulating his social interactive theory of writing, Nystrand criticized his own work for not acknowledging that he was part of the shared social reality of creating meaning "we were acting simultaneously as readers and assessors, failing, in effect to note that the meaning or worth of any text results as much from the act of reading as from the text that is read" (Nystrand 1982a p70). I am the content analyst in this research, and the created meaning between me and the writer when I read and analyse email content is not the meaning created between the writer and intended reader. Suchman describes shared understanding of situations as in part due to the conventional meanings of language expressions, and in part due to their indication of a relation to the circumstances of an actual occasion (Suchman 1987 p58). This point highlights the importance in email content analysis of developing robust and reliable markers, which have predominantly stable shared meanings. My research contributes to this aim and I discuss areas for future research in chapter 9.

Researchers also warn of the risks of not taking level issues into account in research (Gallivan and Benbunan-Fich 2005; Walczuch and Watson 2001). Level issues concern whether data collected at the individual or group level is analysed at the individual or group level. In this research I collected emails as my observations and studied their style by *Sender, Receiver etc.* The communication behaviours analysed are essentially those written in individual emails and I use data from these emails to interpret what is happening at a group level. My research could therefore be criticized for not paying attention to level issues. My defence is the interactivity necessary in team work, and the iterative nature of discourse and social behaviour. The group behaviour is the combined interactions of the individuals' behaviours, which in networked teams is the communication behaviour and the discourse represented in the individual emails.

Regarding developing a standard methodology, I have analysed the data in a systematic way, although I adapted my approach partly according to the focus of each study, partly to avoid losing the relationship with my participants, and partly to what was feasible in each context. Strauss writes:

> *...a standardization of methods (swallowed whole, taken seriously) would only constrain and even stifle social researchers' best efforts... researchers need to be alive not only to the constraints and challenges of research settings and research aims, but to the nature of their data (Strauss 1987 p7).*

4. Pilot study: H1 Influences on email behaviour

4.1 Research focus

This pilot study explores interpersonal email communications during a professional networked team writing project. It is the first study in a series of three analyses to support the development of an email analysis tool, which allows non-intrusive research into writing practice and comparison of different writing projects in a consistent way.

Emails are a form of written communication. Models of writing processes therefore apply both to the document being written in a networked team, and also to the emails written during the project. Influences on writing should therefore influence communication behaviours in the emails, and adaptations identified in communication behaviour may help to predict team writing performance.

The case study reported in this chapter explores email content and dynamics to extract indicators representing influences on writing, and communication behaviours representing the social dimension in networked team writing. Dependencies between the writing influences and communication behaviours are searched for.

> *Pilot study:*
>
> *H1 = Email communication behaviour is the product of writing influences and representative variables of both can be derived non-intrusively from email content.*

Figure 4-1 shows the overall research framework reported in this thesis. The green shaded area (H1) shows the focus of the part reported in this section. This is a secondary analysis of email data collected from a software documentation project in a professional writing context. A preliminary analysis of this data completed within this research framework has been published elsewhere (Edwards et al. 2005). 295 emails were analysed in this study and details of the context and project are presented in section 3.4.2. The content analysis was completed manually from paper copies of the emails and the variables recorded in Excel files for further analysis. Please see section 3.7 for an explanation of the methods used to analyse email content. Writing influences extracted from the email content were *Sender*, *Receiver*, *Send role, Receiver role*, *Direction*, *Distance*, *Purpose* and *Phase*. Variables extracted from email content to represent communication behaviour were elaboration in terms of total word count and % greeting word count, greeting style, and % first person singular and plural pronouns. These markers were used to interpret effort and value attributed to communications, formality, involvement and solidarity. Empirical data derived from the coding were analysed using SAS statistical analysis package.

Closing the loop in terms of testing interpretations through triangulation with qualitative data collected from participants is not possible for this secondary analysis. Therefore, although attempts are made to explain the data, explanations cannot be validated within the scope of this

research. This phase of the research aims to operationalize indicators (as explained in 3.7) and explore whether dependencies emerge. Post analysis qualitative research on projects reported in chapters 5 and 6 will explore participant perceptions to explain and validate the representations of indicators from email data.

In section 4.2, I present the results of the data analysis. Significant results for each marker are then analysed in section 4.3.

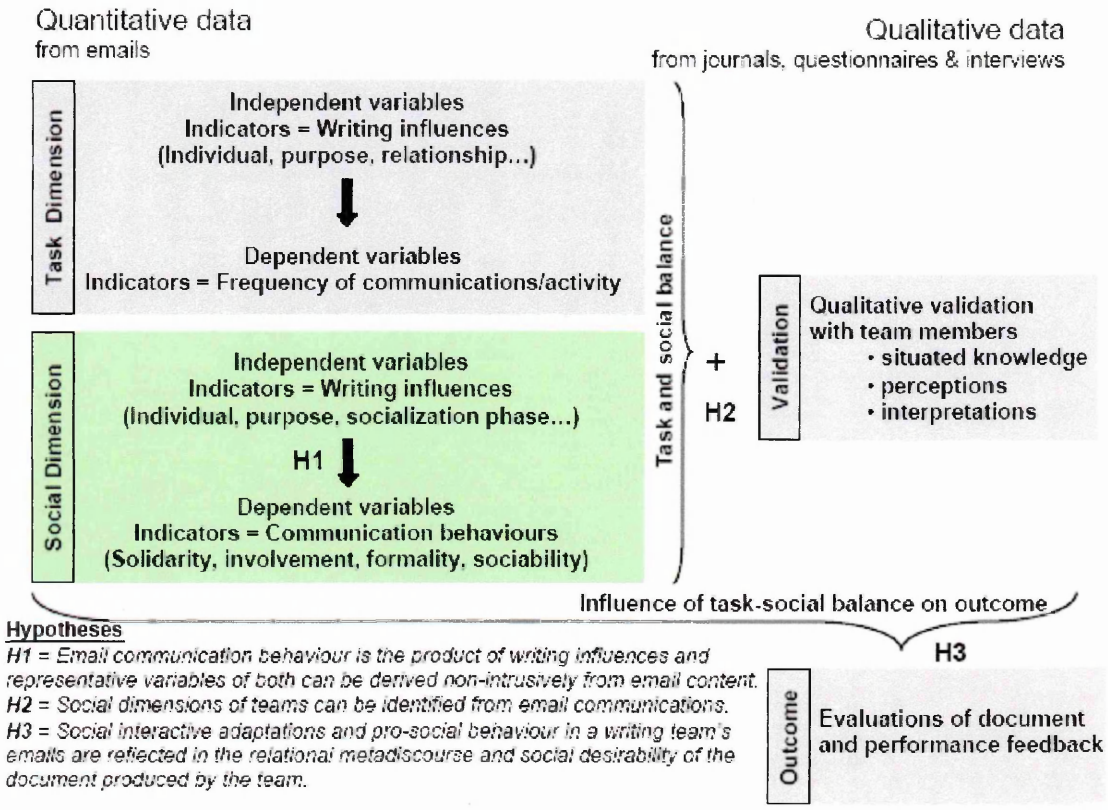

Figure 4-1: Research Framework highlighting H1 phase

4.2 Results: significant writing influences

The Kruskal-Wallis computation tests the null hypothesis that categorical variables (such as *Sender* in this study) have identical distribution functions against the alternative hypothesis that at least two differ with respect to the location. If the p value is small, we can reject the idea that the difference in the location of the groups is a coincidence, and conclude instead that the populations vary. The appropriateness of this test for the data to be analysed is discussed in section 3.7.3. An extract from the SAS output for the Kruskal-Wallis test for the independent variable, *Direction*, and dependent variable, Greeting style, is presented in appendix GG.

Table 4-1 shows the p values for the Chi square value resulting from the Kruskal-Wallis tests for each paired combination of independent and dependent variable, i.e. for each paired combination of a writing influence and email communication behaviour.

Table 4-1: p values for Kruskal-Wallis Chi square values

(Shaded cells indicate p is significant at the 0.05 level)

Pr > Chi square		Dependent variables: communication markers				
		Word count	% Greeting word count	Greeting style / formality score	% first person singular pronouns	% first person plural pronouns
Independent variables: writing influences	Send role	0.0763	0.0131	<0.0001	0.0447	<0.0001
	Receive role	0.0861	0.3683	<0.0001	0.4297	0.0361
	Send level	0.1723	0.0974	<0.0001	0.0481	<0.0001
	Receive level	0.0078	0.8572	<0.0001	0.3655	0.9309
	Direction	0.0062	0.3382	<0.0001	0.1021	0.0041
	Distance	0.2876	0.8731	0.1000	0.4197	0.0061
	Purpose	<0.0001	0.0007	0.0038	0.0005	0.0107
	Phase	0.0773	0.3827	0.3041	0.0765	0.4417

All the writing influences, except *Phase*, caused variations in the communication behaviours in at least one of the markers used. The most prevalent influence was *Purpose*, which affected all the communication behaviours; the communication marker, greeting style, which represented formality, varied with all the writing influences except *Distance* and *Phase.* The marker of solidarity (% first person plural pronouns) varied with all the writing influences, except *Receive level* and *Phase*. Word count and greeting length elaborations are not analysed in depth in this study (explained in section 3.7.2). My discussion of the results on this case study is therefore restricted to interpretations from the formality (greeting style), involvement (first person singular) and solidarity markers (first person plural markers), for the significant results shown in Table 4-1.

The Kruskal-Wallis test can only test whether there are overall differences between the predictor variable categories, but not the nature of the differences (Greene and D'Oliveira 2006 p79). To interpret the nature of the differences between the categories requires interpretation of the trends shown by descriptive statistics of the communication behaviour variables. As mentioned at the beginning of this chapter, it is not possible with this data to validate any interpretations. However, to provide examples of potential interpretations from these email communication behaviour markers, I discuss example results in the following sections.

As some of the markers used in this research have small scales, such as formality, which varies between 1 and 4, and do not have normal distributions, (see section 3.7.3), I am using several descriptive statistics to interpret results. For the formality score, I use the modes, if they vary, to interpret the most frequent behaviour within a category, and for the larger scales for % involvement (0-33) and solidarity (0-13) I consider the means, medians and ranges. However,

regardless of which statistics are most appropriate for interpretations, I have presented all the graphs consistently to present all available location statistics.

4.3 Interpretations of influences

4.3.1 Purpose influence

The effect of *Purpose* was significant on the formality, involvement and solidarity markers (see Figure 4-2 to Figure 4-4). Figure 4-2 shows, for example, that formality in communications was higher when a draft document was being sent for review, or when a request for information was being made, than for the other *Purpose* categories. Figure 4-3 shows that involvement was higher for management and courtesy communications than for other *Purposes*, and Figure 4-4 shows a higher representation of solidarity in communications over product design discussions than in the other *Purposes*. Participants thus adapted levels of formality, and represented themselves and feelings of solidarity with communicating partners differently in the text, according to the *Purpose* of the email. These indicators show, therefore, that communication behaviour varied with the *Purpose* of the communication in this study.

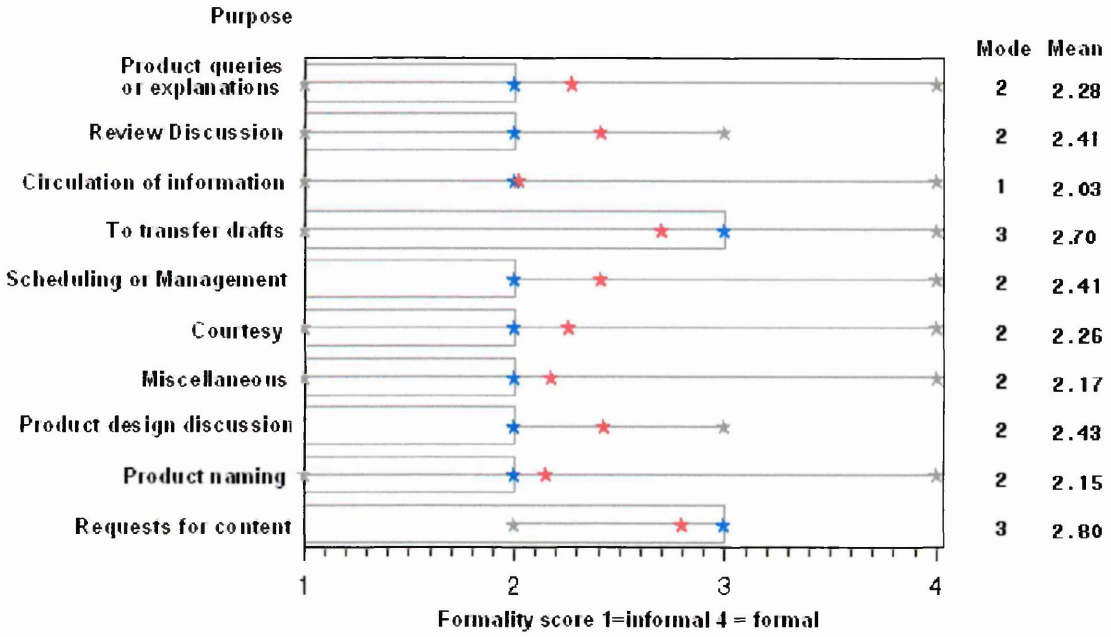

Figure 4-2: Differences in formality by purpose

Legend: min ★— max —★ median ★ mean ★ mode □

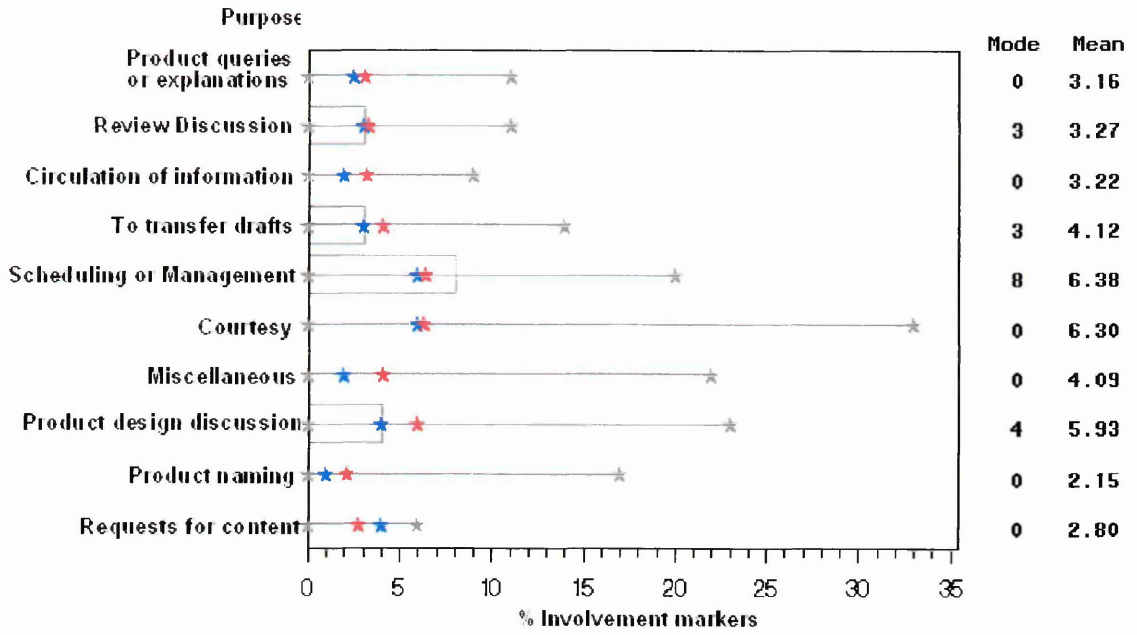

Figure 4-3: Differences in involvement by purpose

Legend: min max median mean mode

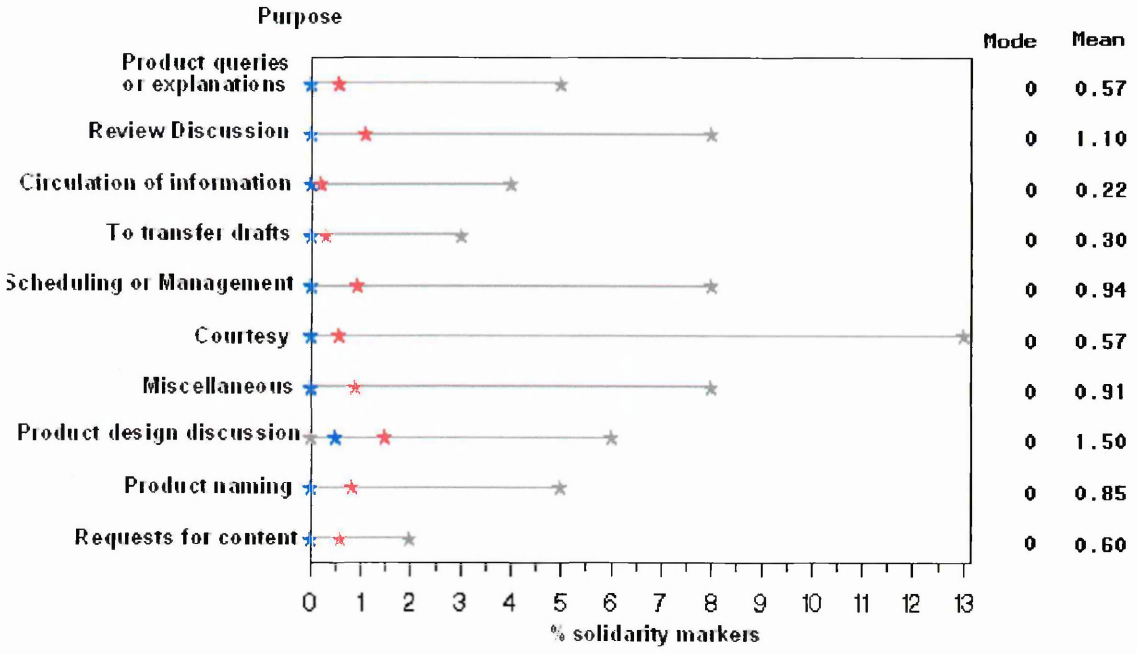

Figure 4-4: Differences in solidarity by purpose

Legend: min max median mean mode

4.3.2 Direction influence

The *Direction* variable describes the transmission direction of the email in the organizational hierarchy. Levels from the company organigram were collapsed into four levels: Vice President (1), Director (2), Manager (3) and Non-manager (4), to which each team member was assigned.

Using the levels for *Senders* and *Receivers*, I coded each email for the direction of transmission in the organizational hierarchy. The effect of *Direction* was significant on the formality and solidarity markers. Figure 4-5 shows that emails sent upwards in the organizational hierarchy were more formal than those sent to an individual of the same level or below and Figure 4-6 shows higher solidarity was expressed in communications downwards in the hierarchy. Thus two of the communication behaviour markers suggest that individuals alter their style of email communication according to the level of the recipient in the organizational hierarchy.

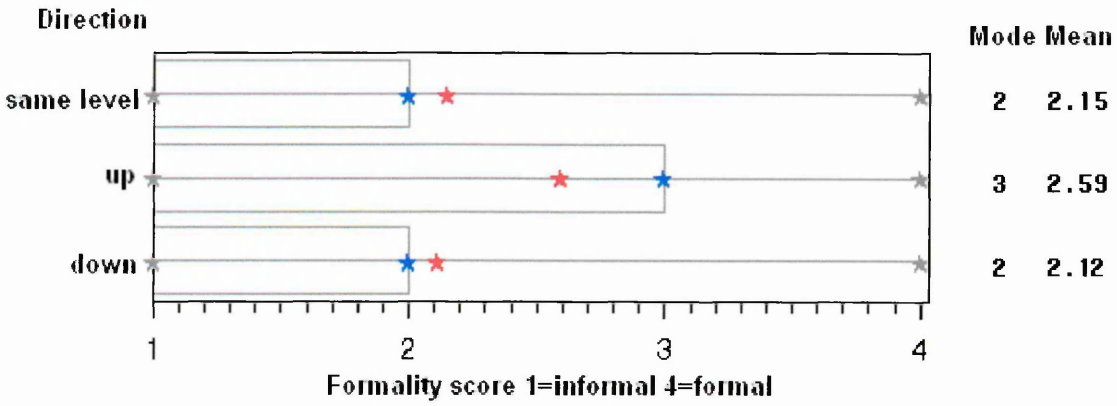

Figure 4-5: Differences in formality by direction in organizational hierarchy

Legend: min max median mean mode

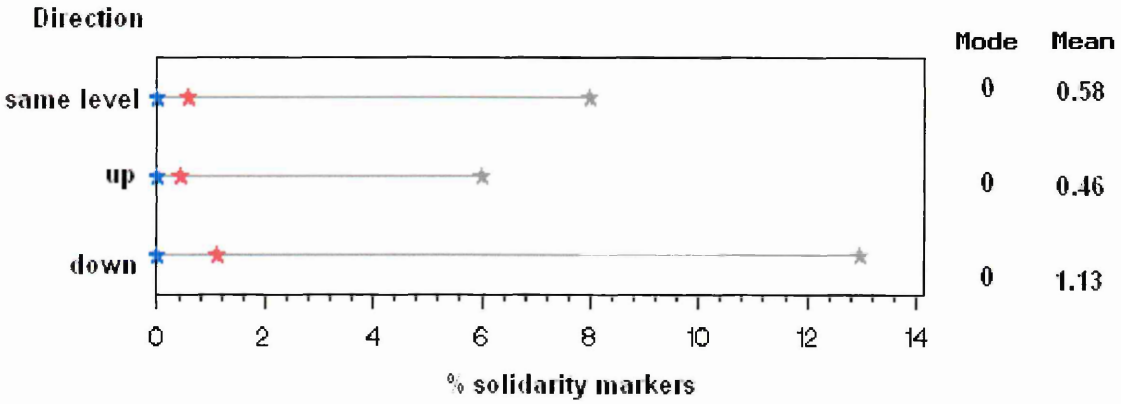

Figure 4-6: Differences in solidarity by direction in organizational hierarchy

Legend: min max median mean mode

4.3.3 Distance

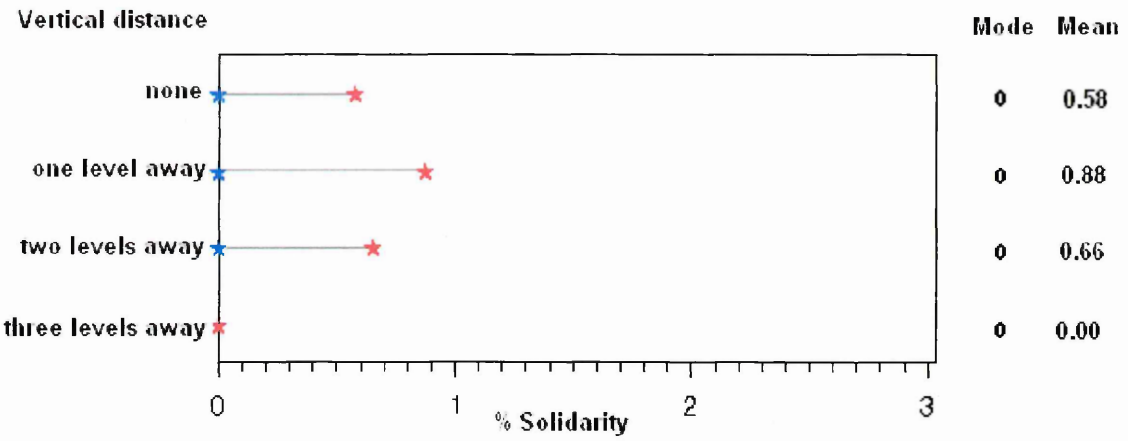

Figure 4-7: Differences in solidarity with vertical distance of transmission

Legend: min ★— max —★ median ★ mean ★ mode □

Vertical distance describes the number of levels in the organizational hierarchy across which an email was sent. Figure 4-7 shows the rather interesting result that expressions of solidarity were higher in emails sent to one level above or below, than they were in emails sent to individuals within the same level. This suggests that individuals felt more solidarity with Managers or subordinates, than they did with colleagues working at the same level. Solidarity decreases as emails are sent further away, i.e. through one, two and three levels.

4.3.4 Receiver level and role

In this data, usually only one person represented each role type, except where team membership changed, and a new team member took over the role of the previous member. This means that interpretations for the *Send* and *Receive roles* can be interpreted as send and receive individuals.

Conforming to the results above related to formality and direction of email transmission in the organizational hierarchy, formality for emails varied with the level of the recipient. Emails sent to the lowest of four organizational levels were less formal than those sent to the other three levels (see Figure 4-8). The individual *Receiver* also influenced the formality of the emails; variations by *Receiver* are shown in Figure 4-9.

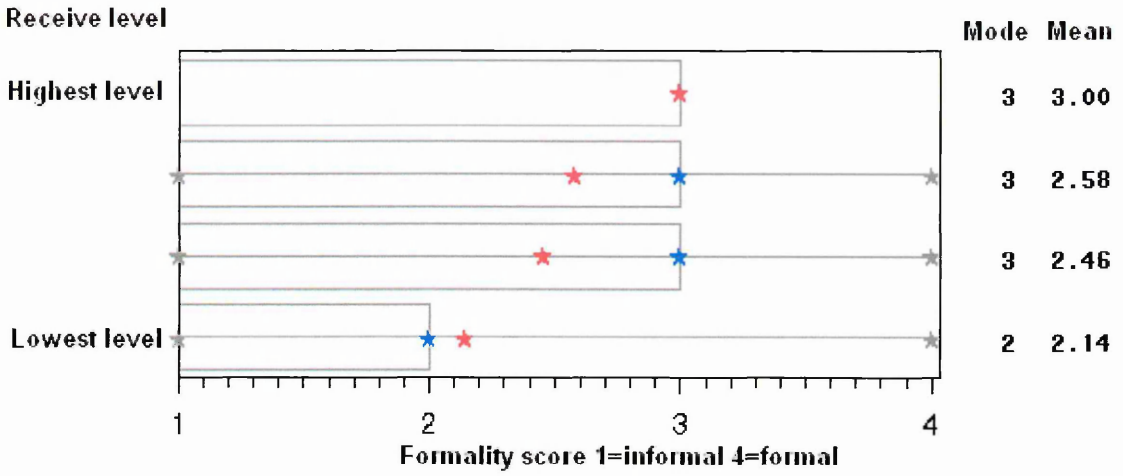

Figure 4-8: Differences in formality by receiver level in organizational hierarchy

Legend: min ★ max ★ median ★ mean ★ mode

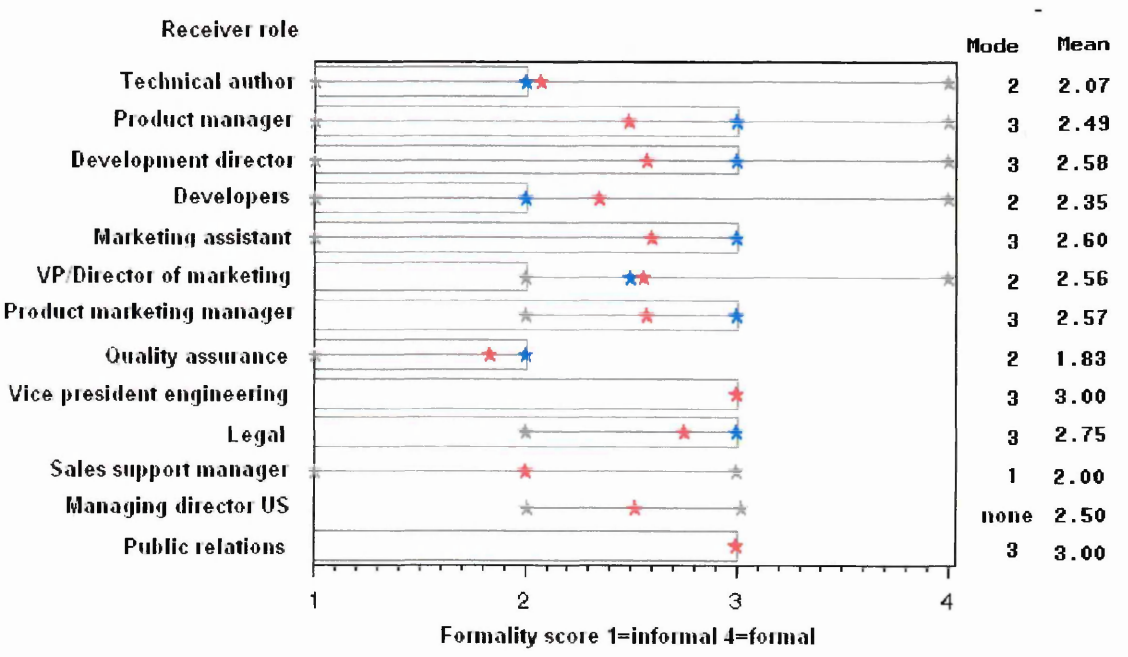

Figure 4-9: Differences in formality by receiver role

Legend: min ★ max ★ median ★ mean ★ mode

Figure 4-10 shows differences in solidarity by *Receiver*. Emails written to the Product Marketing Manager showed the highest solidarity.

These differences by *Receiver* show that writers of emails vary their email formality and representations of solidarity in anticipation of the reader and the reader's level in the organization.

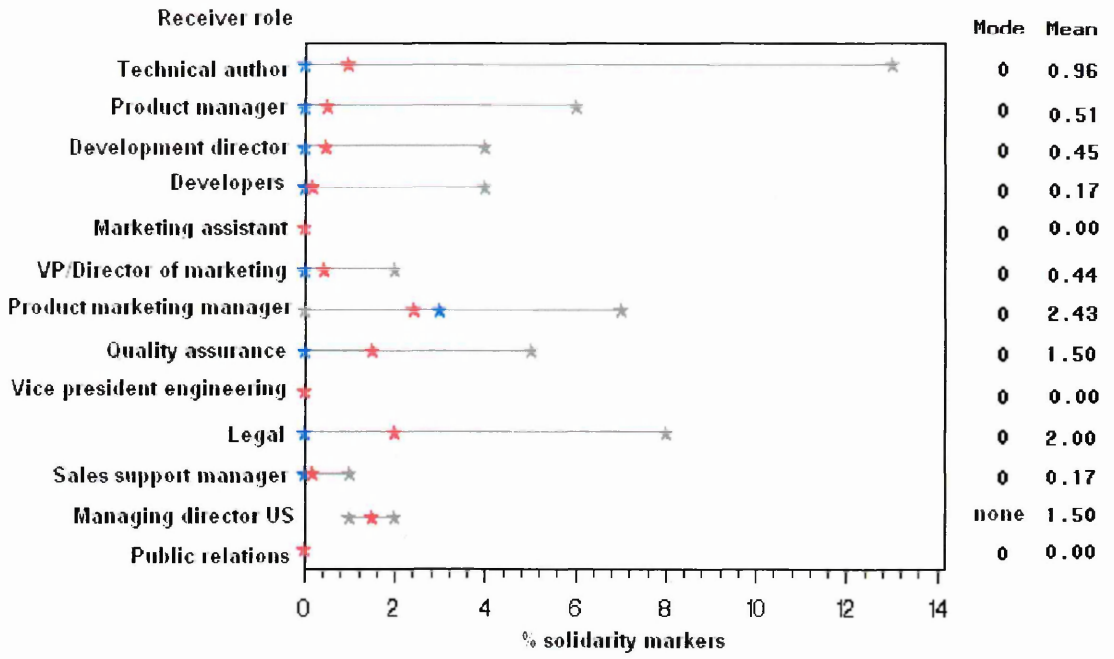

Figure 4-10: Differences in solidarity by receiver role

Legend: min ★— max —★ median ★ mean ★ mode □

4.3.5 Sender level and role

Above in Figure 4-5 and Figure 4-8, we have seen that formality is highest in emails sent upwards in the organizational hierarchy and lowest in emails sent to the lowest level in the hierarchy respectively. Here Figure 4-11 shows that emails written by the members of the lowest level in the hierarchy are the most formal. Thus, emails *written* by the lowest level members are most formal, emails *received* by the lowest level members are least formal, and emails sent *upwards* are most formal. Analysis of the social markers from these emails therefore informs on the communication and social norms, which in this case reflect the power relationship in the organizational hierarchy.

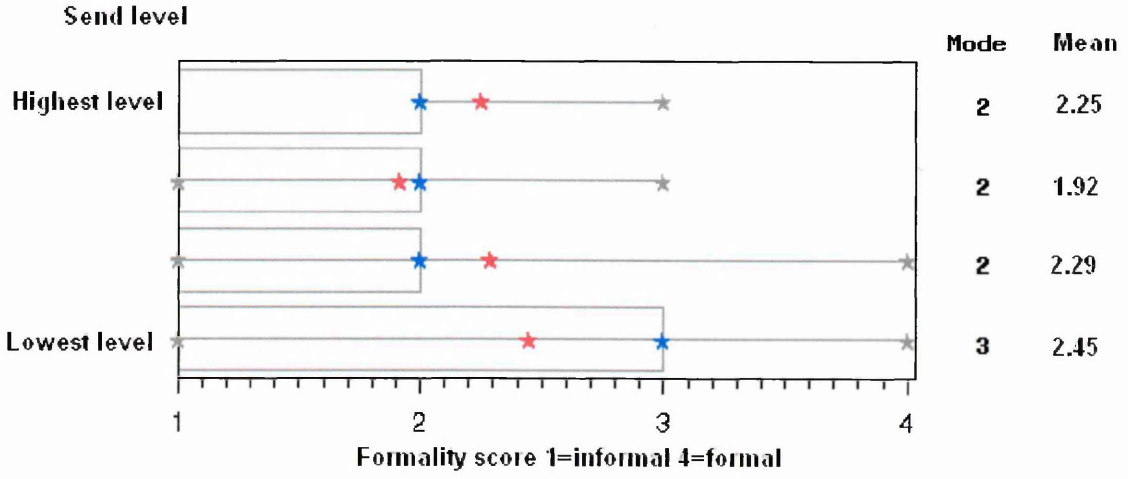

Figure 4-11: Differences in formality by sender's level

Legend: min ★— max —★ median ★ mean ★ mode □

The *Sender's level* in the organization also influenced representations of involvement and solidarity in emails (see Figure 4-12 and Figure 4-13). The highest level members show the highest involvement and the two middle levels show relatively higher solidarity than the two extreme levels.

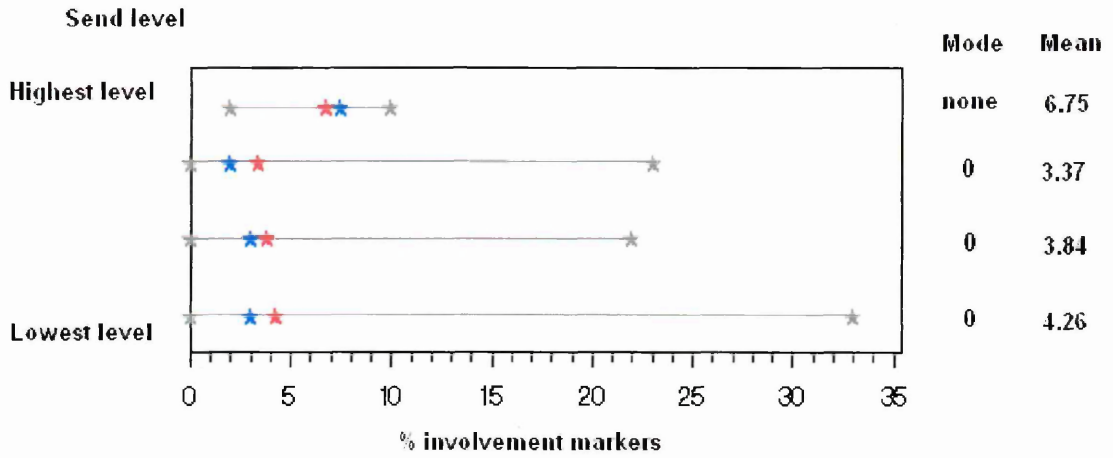

Figure 4-12: Differences in involvement by sender's level

Legend: min ★— max —★ median ★ mean ★ mode □

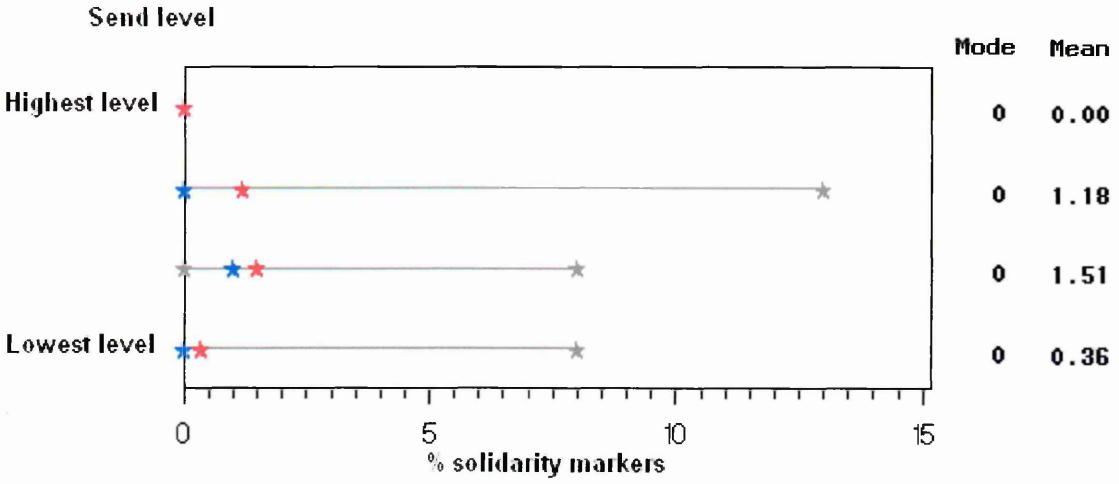

Figure 4-13: Differences in solidarity by sender's level

Legend: min ★— max —★ median ★ mean ★ mode □

Further, *Send role*, or individual, influenced formality, involvement and solidarity. For example, Figure 4-14 shows that the Technical Author had the most formal email style; the Project Manager shows relatively high involvement in Figure 4-15 and the Sales Support Manager shows relatively high solidarity in Figure 4-16. Thus individual writers varied in their style of email communication, as represented by these three communication markers.

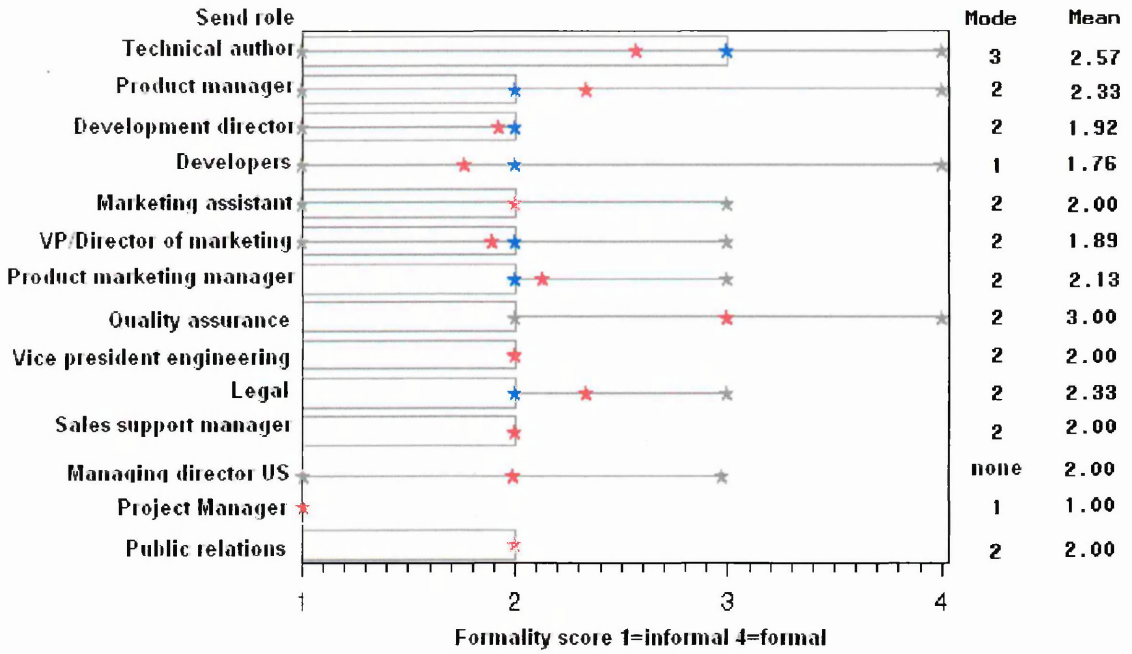

Figure 4-14: Differences in formality by sender's role

Legend: min ★— max —★ median ★ mean ★ mode □

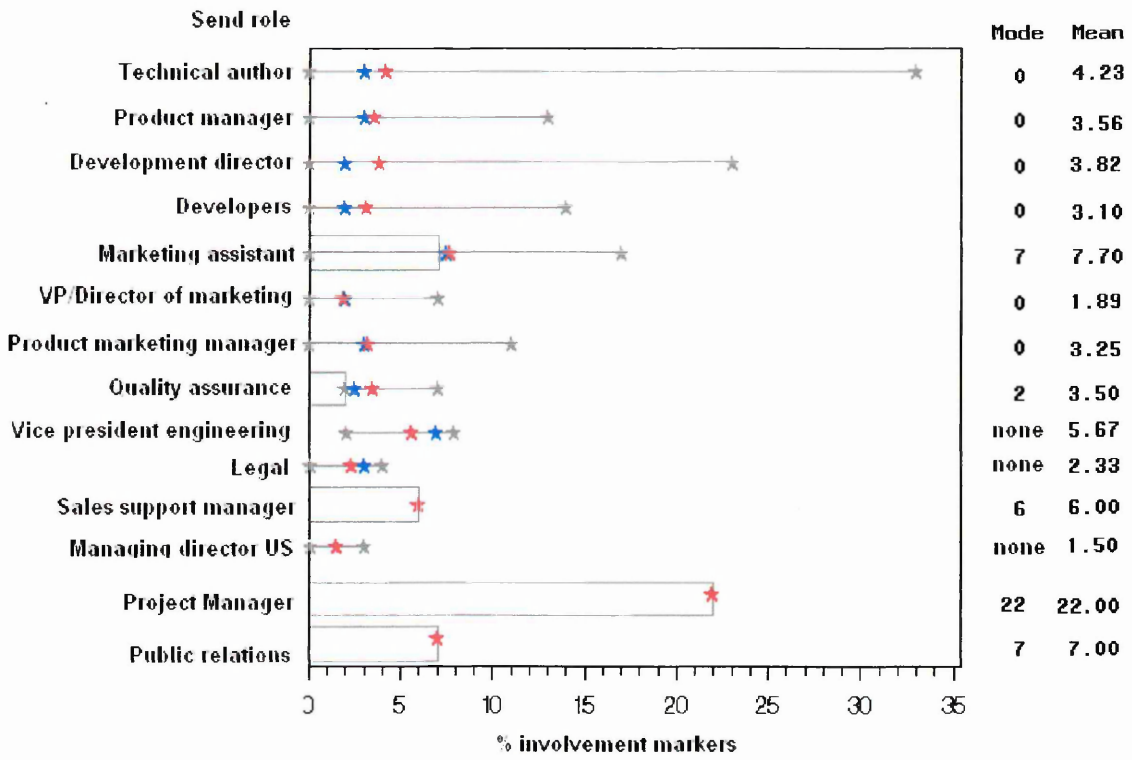

Figure 4-15: Differences in involvement by sender's role

Legend: min ★— max —★ median ★ mean ★ mode □

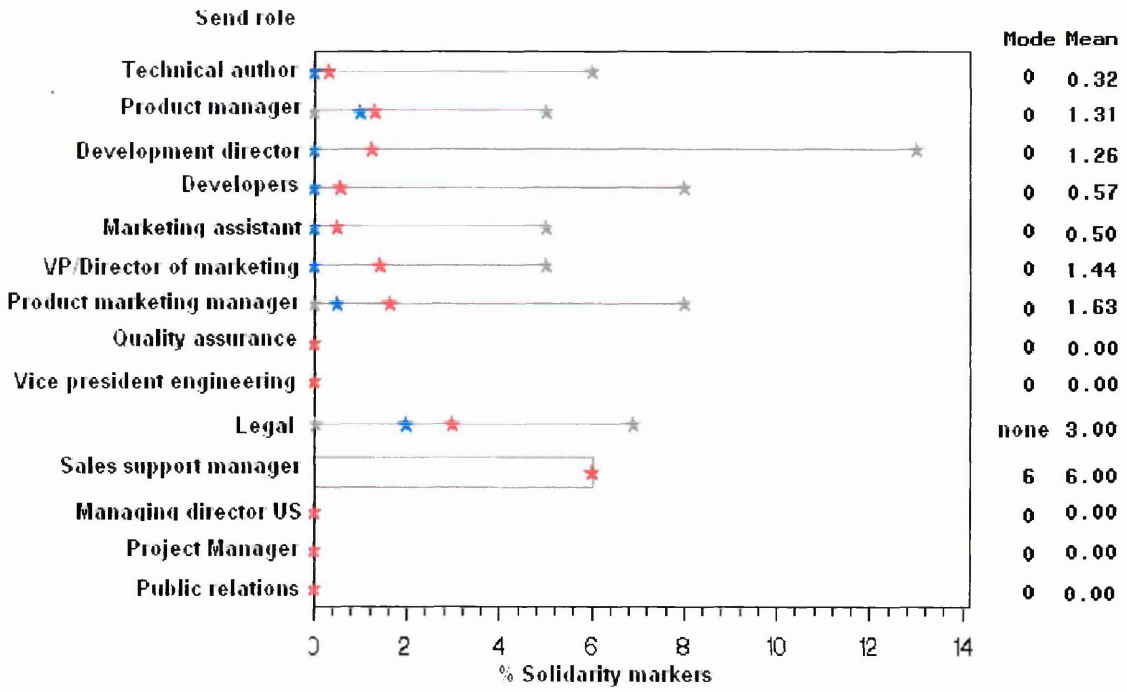

Figure 4-16: Differences in solidarity by sender's role

Legend: min ★— max —★ median ★ mean ★ mode □

4.4 Discussion

All the writing influences, except *Phase*, caused variations in the communication behaviours in at least one of the markers used. The most prevalent influence was *Purpose*, which affected all the communication behaviours; the communication marker, greeting style, which represented formality, varied with all the writing influences except *Distance* and *Phase.* The marker of solidarity (% first person plural pronouns) varied with all the writing influences, except *Receive level* and *Phase*. Although the results are significant in terms of rejecting the null hypotheses that there is no variation in email communication behaviours with the writing influences, this is only a small case study analysing 295 emails from 14 people on one writing project. Nevertheless a cautious view suggests there is value in using the methods developed here to explore communication variables and predictors in writing processes.

Studying trends in distributions of communication behaviour markers across the different writing influence categories has the potential to describe the social dimension of networked team writing.

There were significant differences with *Purpose* for formality, involvement and solidarity. Formality was higher in emails being sent for review or when a request for information was being made, than for other purpose categories. Involvement was highest for management and courtesy communications, and higher solidarity was expressed in emails discussing product design than for other purposes.

Formality and solidarity varied with *Direction*; formality was higher in emails sent upwards in the organizational hierarchy than in those sent to individuals at the same level or below, and solidarity was highest in emails sent downwards in the hierarchy.

Emails sent to *Receivers* in the lowest of four organizational levels were less formal than those sent to the other three levels. Differences in formality and solidarity by *Receiver* show that writers adapt their email style for their intended readers and the readers' level in the organization.

Combining interpretations from the markers helps to develop a profile of the social behaviour in the team. Emails *written* by the lowest level members were most formal, emails *received* by the lowest members were least formal, and emails sent upwards were most formal. Analysis of these markers thus informs on the communication and social norms, which in this case reflect the power relationship in the organizational hierarchy.

Solidarity also varied with *Distance*: individuals expressed more solidarity with Managers or subordinates, than they did with colleagues working at the same level. Highest level members show the highest involvement and the two middle levels show relatively higher solidarity than the two extreme levels. Thus the middle-management showed more solidarity than the top management or non-managers. Formality, involvement and solidarity also differed individually.

Three benefits will emerge from the use of such representations:

- The data can be compared with other qualitative data collected from interviews or questionnaires. For example, Figure 4-9 shows that writers vary their level of formality in anticipation of the recipient. A questionnaire asking respondents to rate likeability of other partners may show that these communication markers reflect the closeness of relationships in teams and thus describe the social dimension of the team.
- The kind of analysis provided in this study may be useful within ongoing projects as an intervention tool. Interpreting the data together with situated knowledge from participants may help to improve the social dimension of the project. For example, in this project, I was the Technical Author. In Figure 4-14, you can see that I was the most formal team member on this project and this has led me to re-evaluate my own professional communication style in emails.
- Comparisons can be made across writing groups in different contexts, allowing a meta analysis. Thus a standard methodology can be applied to multiple projects without researcher interference in the context of work. Results of cross context studies using standard methodologies are more likely to produce broadly applicable results.
- Certain behaviours or combinations of behaviours may result in more successful projects. Studying the process and outcomes of team writing groups may thus help towards developing a causal model for team writing and identifying predictors of successful outcomes.

Although these results may not be transferable to other writing contexts, they indicate that communication behaviour in emails appears to vary with influences on the writing process, and that this can be shown quantitatively by extracting variables from email content.

Indicators which represent communication behaviour or writing influences in this particular study need to be tested for validity as follows:

- Using different samples of email communications from academic and professional writing contexts. Corroborating data from different projects and contexts will build more convincing evidence of validity.
- By surveying participants within contexts post analysis to reinforce interpretations from the email data with writers' situated knowledge and perceptions.

4.5 Chapter review

This study focused on the hypothesis

> *Pilot study:*
>
> *H1 = Email communication behaviour is the product of writing influences and representative variables of both can be derived non-intrusively from email content.*

In this study, communication behaviour was represented by five dependent variables extracted from email content: elaboration in terms of total word count and % greeting word count, greeting style and % first person singular and plural pronouns. These markers were used to interpret effort and value attributed to communications, formality, involvement and solidarity. Writing influences also extracted from the email content were *Sender*, *Receiver*, *Send role, Receiver role*, *Direction*, *Distance*, *Purpose* and *Phase*.

All the writing influences, except *Phase*, caused variations in the communication behaviours in at least one of the communication markers used. The most prevalent influence was *Purpose*, which affected all the communication behaviours; the formality communication marker was influenced by all the writing influences except *Distance* and *Phase* and the solidarity marker was affected by all the writing influences, except *Receive level* and *Phase*.

Greeting style, involvement and solidarity were studied in further depth, for their potential to describe the social dimension of networked writing teams. Interpretations from the data helped to profile the social dimension of team working, for example by identifying formality norms which mirror the organizational hierarchy, and vertical solidarity in the organization rather than horizontal solidarity.

These results indicate that communication behaviour in emails varies with writing influences and this can be shown quantitatively within context, by extracting variables from email content. Interpretations are possible from the adaptations in communication behaviour, although validation requires situated knowledge from the team writing context. This can be achieved by collecting qualitative data from participants post analysis, a technique explored in the next case study

(chapter 5). The method is improved on and repeated in two further case studies, using content analysis software, to explore the feasibility of using adaptations in email communication behaviour to understand how team culture can influence virtual team writing.

5. H2 Commercial case study: team social dimensions

5.1 Research focus

This study explores interpersonal email communications during a professional networked team writing project in a client-supplier scenario. It is the second study in a series of analyses to support the development of an email analysis tool, which allows non-intrusive research into writing practice and comparison of different writing projects in a consistent way.

The previous study reported in chapter 4 showed that indicators of influences on writing could be derived from email records and shown to influence social behaviour in email communications in a networked team writing project. This study applies the same process to test whether dependencies occur in a second commercial context, and to interpret these in the light of participant feedback. This part of the research is shaded green in Figure 5-1, which shows how this study fits within the overall research framework and addresses the following hypothesis:

H2 = Social dimensions of teams can be identified from email communications.

218 emails from the technical writing project were coded using MAXqda content analysis software. Writing influences analysed were the writer, receiver, audience size, relational direction, purpose of the email, socialisation phase and language. The communication markers used to interpret pro-social behaviour in team work were elaboration in body text and greetings of emails, style of greetings, use of first person pronouns and social building units. These markers were used to interpret effort and value attributed to communications, formality, involvement, solidarity and sociability.

Empirical data derived from the coding were analysed using SAS statistical analysis package. Pre and post analysis interviews were completed with one representative from each of the client and supplier organizations and the eight main participants were asked to complete a questionnaire. Perception data collected from the team members were used to derive a context-sensitive formality scale for emails on the project, and to test the validity of interpretations. For a full description of the methods used in this analysis and the background to the Namahn-Banksys project, please see chapter 3.

In section 5.2, I present the results of the data analysis. Significant results for each pro-social marker are then analysed in depth in section 5.3 and triangulated with interview and questionnaire data to develop interpretations.

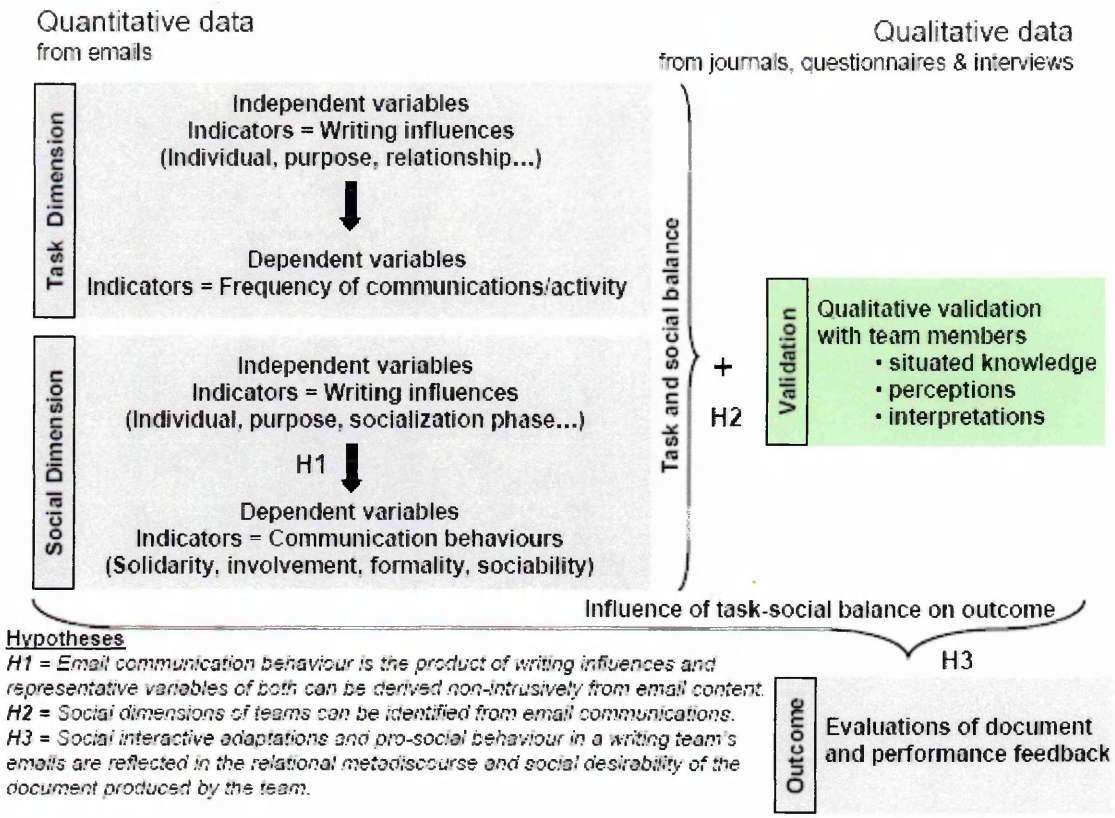

Figure 5-1: Research Framework highlighting H2 phase

5.2 Results: significant writing influences

The Kruskal-Wallis computation tests the null hypothesis that categorical variables (such as *Sender* in this study) have identical distribution functions against the alternative hypothesis that at least two differ with respect to the location. If the p value is small, we can reject the idea that the difference in the location of the groups is a coincidence, and conclude instead that the populations vary. The appropriateness of this test for the data to be analysed is discussed in section 3.7.3. An extract from the SAS output for the Kruskal-Wallis test for the independent variable, *Purpose,* and dependent variable, social building units, is presented in appendix HH.

Table 5-1 shows the p values for the Chi square value resulting from the Kruskal-Wallis tests for each paired combination of independent and dependent variable, i.e. for each paired combination of a writing influence and email communication marker.

Table 5-1: p values for Kruskal-Wallis Chi square values

(Shaded cells indicate p is significant at the 0.05 level)

Pr > Chi square		Dependent variables: communication markers						
		Word count	Open greeting	Close greeting	% first person singular pronouns	% first person plural pronouns	Formality	Social building units
Independent Variables: writing influences	Sender	0.0900	<0.0001	<0.0001	0.0191	0.2609	<0.0001	0.0909
	Receiver	0.0078	0.1381	0.5033	0.1165	0.2098	0.0016	0.0118
	Audience	0.0122	0.4392	0.3416	0.0029	0.0104	0.3883	0.0037
	Direction	0.0156	<0.0001	<0.0001	0.0103	0.0826	<0.0001	0.0507
	Purpose	<0.0001	0.0025	0.0149	0.0001	0.0309	0.0046	<0.0001
	Phase	0.0422	0.0293	0.1563	0.6397	0.0147	0.0005	0.2144
	Language	0.4727	0.0498	0.4014	0.1108	0.2745	0.9314	0.0172

All the writing influences studied caused variations in the communication behaviours in at least three of the markers used. For each of the gray shaded cells in Table 5-1, $p \leq 0.05$. For example, for formality, the p value for the Chi square result from the Kruskal-Wallis test for class *Sender*, is <0.0001. We can therefore reject the null hypothesis that there are no overall differences between formality levels measured for different *Senders*. At least two of the email writers in this project varied in their style of email formality as measured by the formality score. The most prevalent influence was *Purpose*, which showed differences in all the communication behaviour markers and which represents the task dimension.

The Kruskal-Wallis test can only test whether there are overall differences between the variable categories, but not the nature of the differences (Greene and D'Oliveira 2006 p79). To interpret the nature of the differences between the categories requires interpretation of the trends shown by descriptive statistics of the communication behaviour variables and these are discussed further in the following sections, together with feedback from the interviewees, where appropriate.

The analyses have been iteratively checked, corrected and rerun since the preliminary analyses presented to the interviewees, so that some of the results presented in appendices N and O do not match the final corrected results reported in this chapter. The questions used in the post analysis interviews, however, were sufficiently open to invite general comments, which still help to develop meaningful interpretations of the trends shown in the data.

As many of the markers used in this research have small scales, such as open greeting length which varies between 0 and 5 words, and do not have normal distributions, (see section 3.7.3), I am using several descriptive statistics to interpret results. Where the scales are small, I use the modes, if they vary, to interpret the most frequent behaviour within a category. For larger scales,

such as % involvement (0-15) or word count (0-503) I consider the means, medians and ranges. However, regardless of which statistics are most appropriate for interpretations, I have presented all the graphs consistently to present all available location statistics.

5.3 Interpretations of social behaviours

5.3.1 Value attributed to communications: elaboration

Word count was used as a marker of pro-social behaviour, with amount of elaboration representing effort and importance attributed to a communication. Elaboration varied with *Receiver*, *Audience size, Direction, Purpose and Phase*. Thus authors of emails adapted the extent to which they elaborated according to the addressee, how many people the email was visible to, their organizational relationship with the receiver, the task and socialisation in the team.

Word count varied between 0 and 503 words. The highest mean word counts *by Receiver* (see Figure 5-2) were for emails written to the client team and to the supplier technical reviewer. The *Receiver* labelled "Client team" included up to six Banksys employees. The high elaboration in these emails therefore reflects the importance attributed by writers to communications written by the supplier and intended for the client (62% emails) or emails written in house to colleagues and superiors by the client (38% emails).

The high elaboration in the emails to the supplier technical reviewer (30 emails) is explained by this team member's involvement initially as the project leader, until phase two.

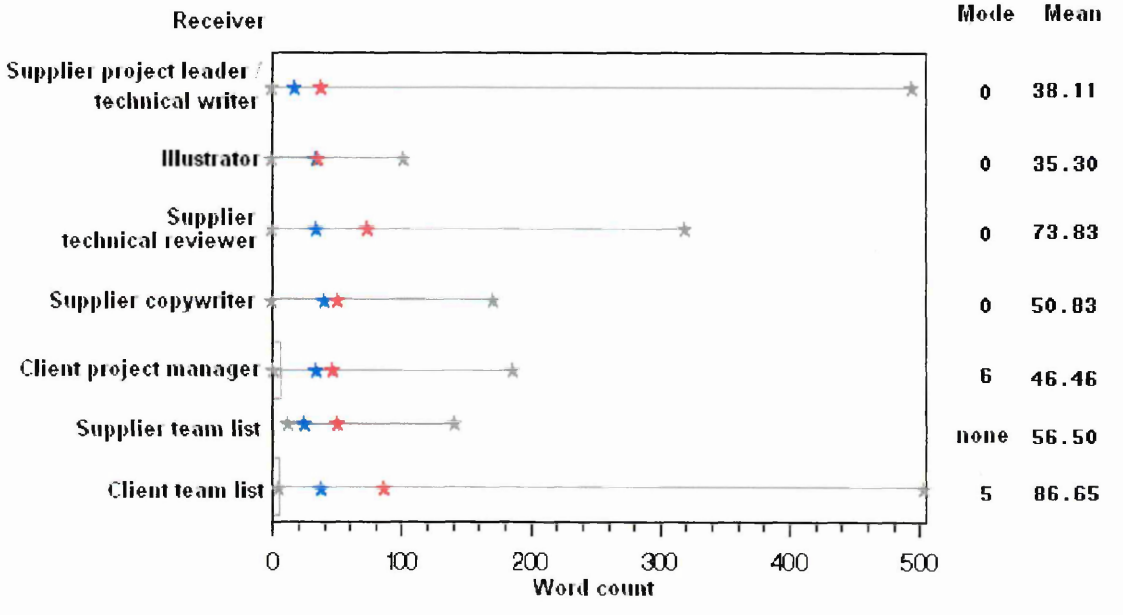

Figure 5-2: Differences by receiver for body text word count

Legend: min ★— max —★ median ★ mean ★ mode

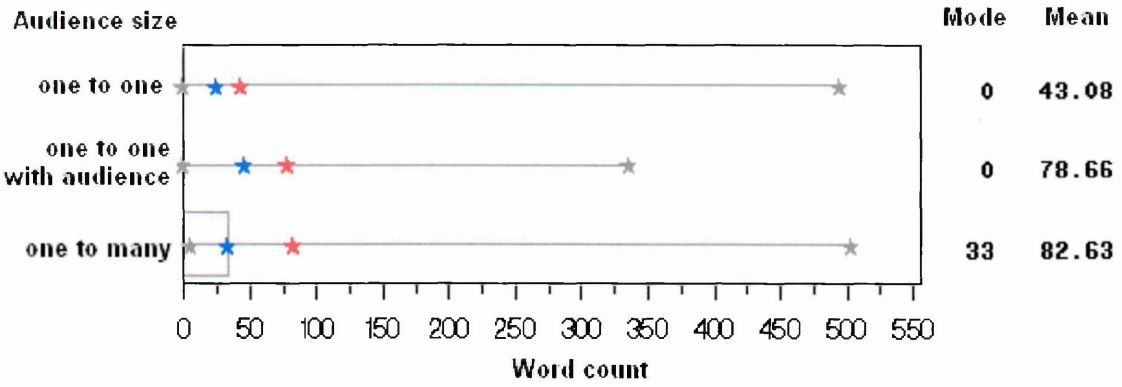

Figure 5-3: Differences by audience size for body text word count

Legend: min ★ max ★ median ★ mean ★ mode □

Figure 5-3 shows the differences between emails intended for a single recipient, those addressed to a single recipient, but also copied to others, and emails intended for multiple recipients. The means increase with increasing audience size. Thus when it became important to write to more than one individual on a topic, this importance was also reflected in the elaboration.

Figure 5-4 shows descriptive statistics for word count by *Purpose*. Tasks with higher means and therefore requiring higher elaboration were *Review/revision and Document design*. *Document transfer* required low elaboration. Thus we see a polarisation between elaboration levels according to whether the communication content represented the task e.g. a review discussion, or served as a transfer agent, e.g. for document transfer.

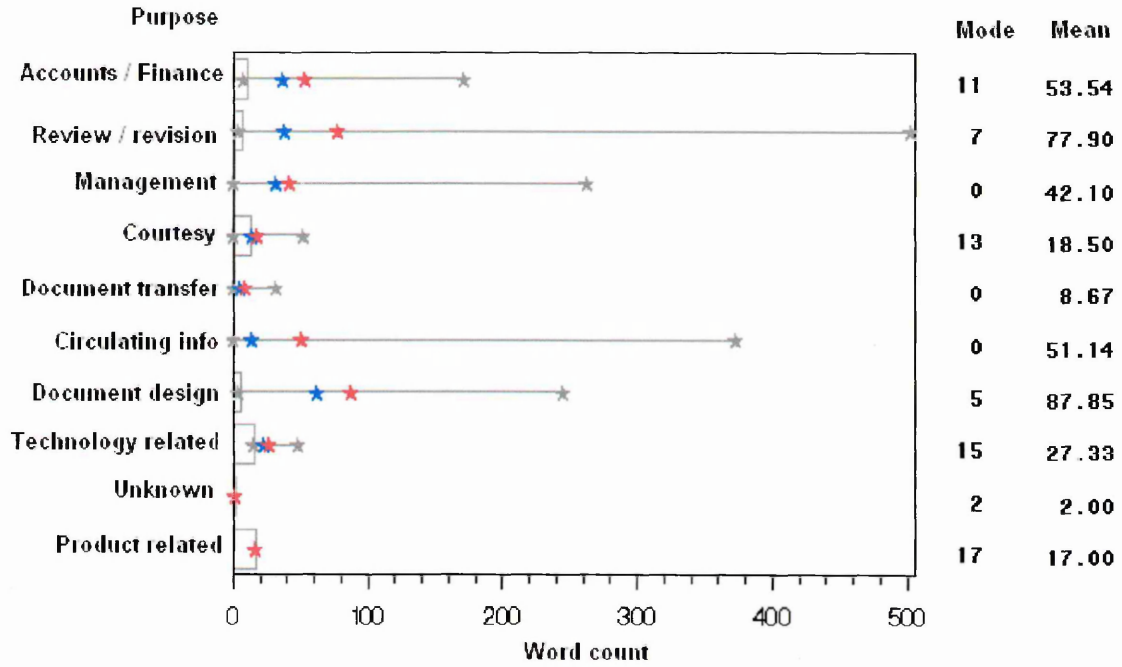

Figure 5-4: Differences by purpose for body text word count

Legend: min ★ max ★ median ★ mean ★ mode □

According to the Kruskal-Wallis test, distributions for word count also varied significantly for *Direction* of emails (p=0.0156); differences are shown in Figure 5-5. The means for word counts show us that client authors elaborated more for in house communications than when writing to the supplier. The reverse was true for the supplier, with more elaboration in communications to the client than for in house communications. The results thus suggest that the supplier makes more effort and attributes more value for communications to the client, than for communications in house, and that the client makes less effort and attributes less value to communications for the supplier than for communications in house. Profiles for elaboration by relational direction thus reflect the power hierarchy of the business relationship.

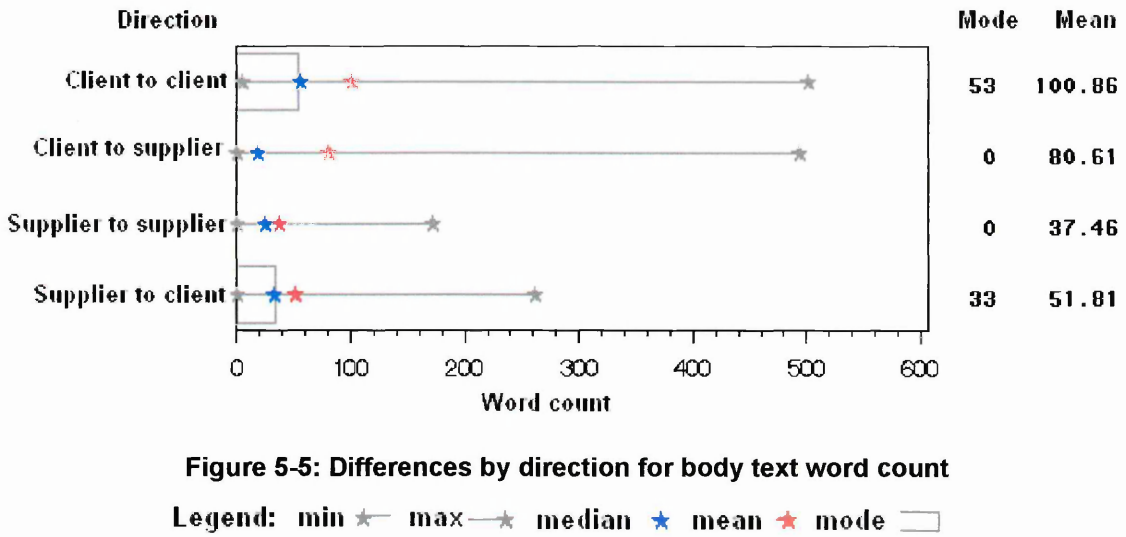

Figure 5-5: Differences by direction for body text word count

The Kruskal-Wallis test also showed significant differences for elaboration between at least two of the socialisation phases of the project. From the means in Figure 5-6, we can see that elaboration peaks at the beginning, middle and end of the project. The project ran from November 2003 until January 2004. Mapping the socialisation phases against the calendar timeline of the project, phases 2 to 5 include the main activities of the project covering calendar dates from 4th November until 14th January. The midpoint of the project falls at the beginning of phase 4, and just a few days after the first of two document review meetings. The final document review meeting was held at the end of phase 5. In terms of Gersick's punctuated equilibrium model (Gersick 1988 p117; see section 2.3.2 on "Development phases"), therefore, higher means for elaboration in communications coincide with the beginning, midpoint transition and end of the project, also the timing of the three FtF meetings.

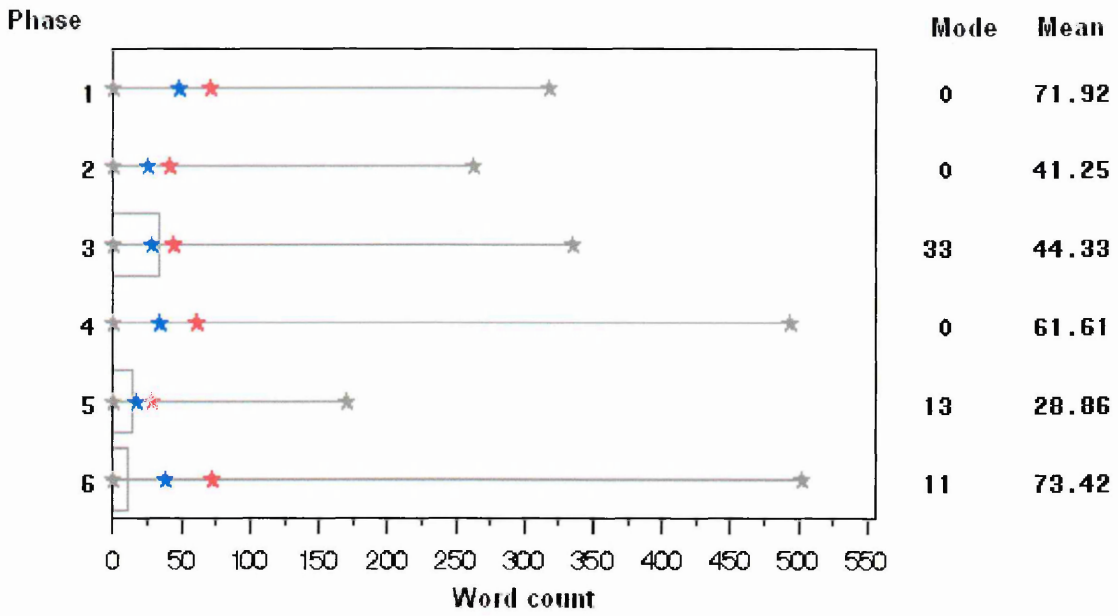

Figure 5-6: Differences by phase for body text word count

Legend: min ★— max —★ median ★ mean ★ mode □

5.3.2 Greeting length as a marker of formality

I have analysed and discussed the social significance of greeting length in communication behaviour in this particular study in section 3.7.2, where I concluded that in this data, increasing greeting length represents increasing formality. Open and close greeting lengths varied with *Sender, Direction* and *Purpose.* Open greeting length additionally showed differences by *Phase* and *Language.* Thus evidence from two markers of formality (in this case study) suggest that writers adapt their style individually, with the relationship between the writer and reader (e.g. client to supplier etc.) and with the purpose of the communication. Further, open greeting lengths suggest differences in formality by *Language* and that writers vary their formality style with socialisation.

Figure 5-7 and Figure 5-8 show the trends for open and close greeting word counts for which there are significant differences between at least two *Senders* ($p < 0.0001$). Here the scale is small and I therefore base my interpretations on the modes. Interpreting open greeting length, as a marker of formality, we see that there are two levels of formality distinguishing the *Senders. Senders* with higher modes for open greeting length, and therefore higher formality style are the four supplier team members, and *Senders* with the lower open greeting lengths and therefore lower formality style are the two client team members.

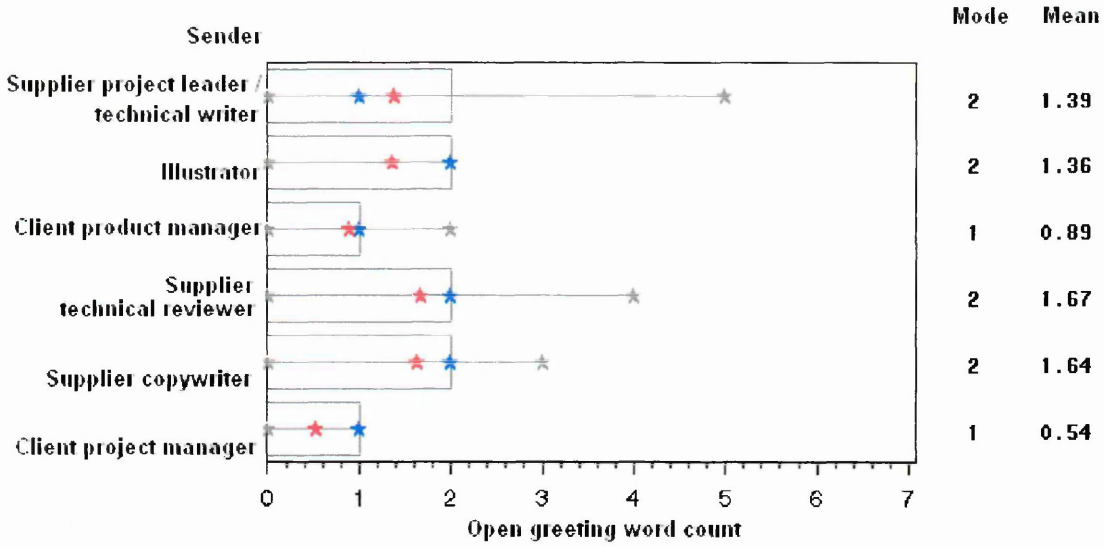

Figure 5-7: Differences by sender for open greeting word count

Legend: min ★ max ★ median ★ mean ★ mode □

The statistical difference in open greeting lengths by *Sender* was reinforced by interview data. In response to the statement that elaboration in open greetings varied by *Sender*, the supplier and client interviewees responded:

> *Yes, of course I would expect this. Everyone uses their own style of communication; this is no surprise (supplier: appendix N, line 8).*
>
> *Yes I expected this. Banksys has no regulations on email style, so everyone starts their emails how they want to. I don't find this strange (client: appendix O line 9)*

Thus communication behaviour in email open greetings is partly individualistic, and the results show higher formality in emails written by team members from the supplier organization than by team members in the client organization.

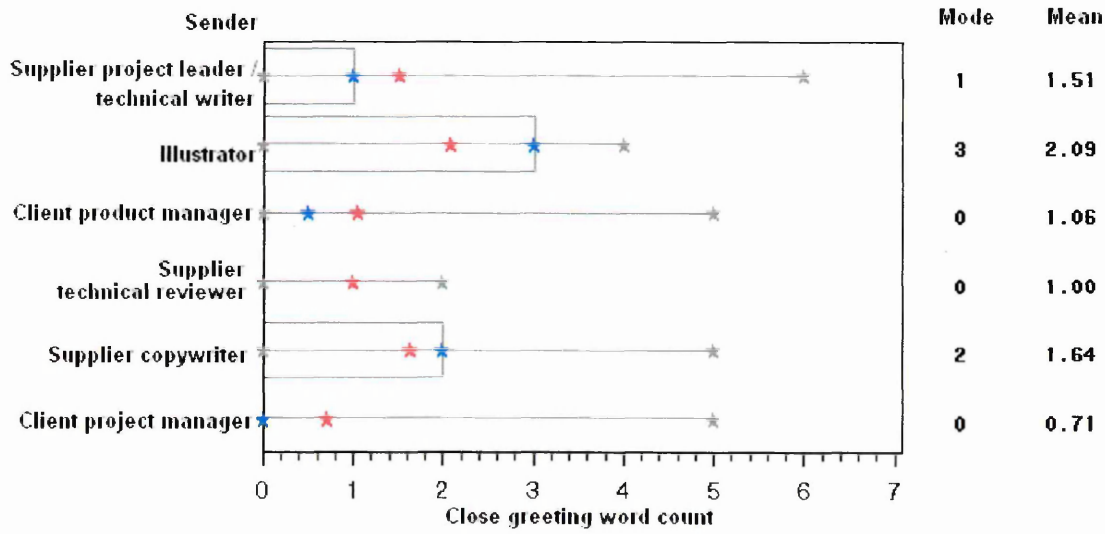

Figure 5-8: Differences by sender for close greeting word count

Legend: min ★ max ★ median ★ mean ★ mode □

Likewise, close greeting varied by *Sender* (see Figure 5-8). When probed to comment on the relatively low (mean) value for the client project manager, the interviewees commented as follows:

> *Yes, of course I would expect this…..he has an informal different style (supplier: appendix N, line 37).*
>
> *I can't think of any particular reason for this. I notice XXX just signs a name sometimes. Perhaps this is the reason (client: appendix O, line 27).*

Thus, style of close greetings is also partly individualistic.

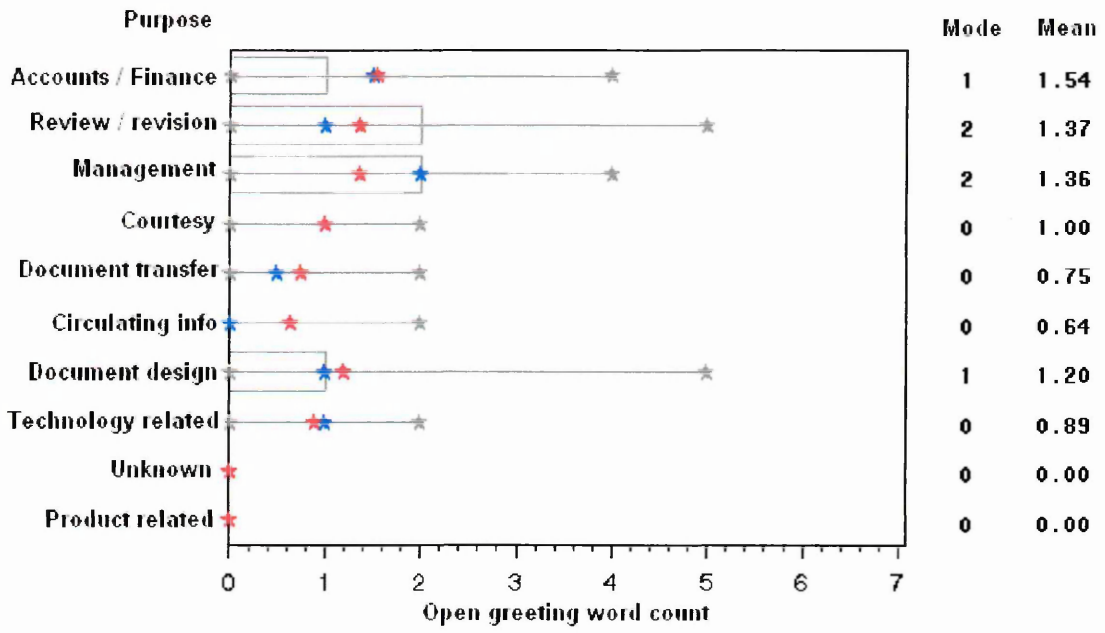

Figure 5-9: Differences by purpose for open greeting word count

Legend: min ★— max —★ median ★ mean ★ mode □

Variation in open and close greeting with *Purpose* suggests that authors adjust their greeting behaviour according to the task at hand. Open greeting modes are highest for *Accounts, Review, Management and Document Design* (see Figure 5-9). Close greeting modes are highest for *Accounts, Review, Management, Courtesy and Product discussions* (see Figure 5-10). Interviewees responded to preliminary results on greeting lengths as follows:

> *For anything to do with accounts we are usually more polite, this is to be expected. We're always more careful in communications on finance or about a review; that's the moment of truth, the review (supplier: appendix N, line 45).*
>
> *I would expect this pattern. At the beginning of the project, when most of these discussions [accounts] happen, people don't know each other and are more official, exploring to see if they can work together (client: appendix O, line 33).*

These interpretations suggest that a longer closing greeting in an email, representing a more formal relationship, is used for certain purposes, either because those purposes are completed early on e.g. *Accounts* and *Document design*, when people do not know each other so well, or for purposes

for which more care is needed and over which people feel more anxious e.g. *Review*.

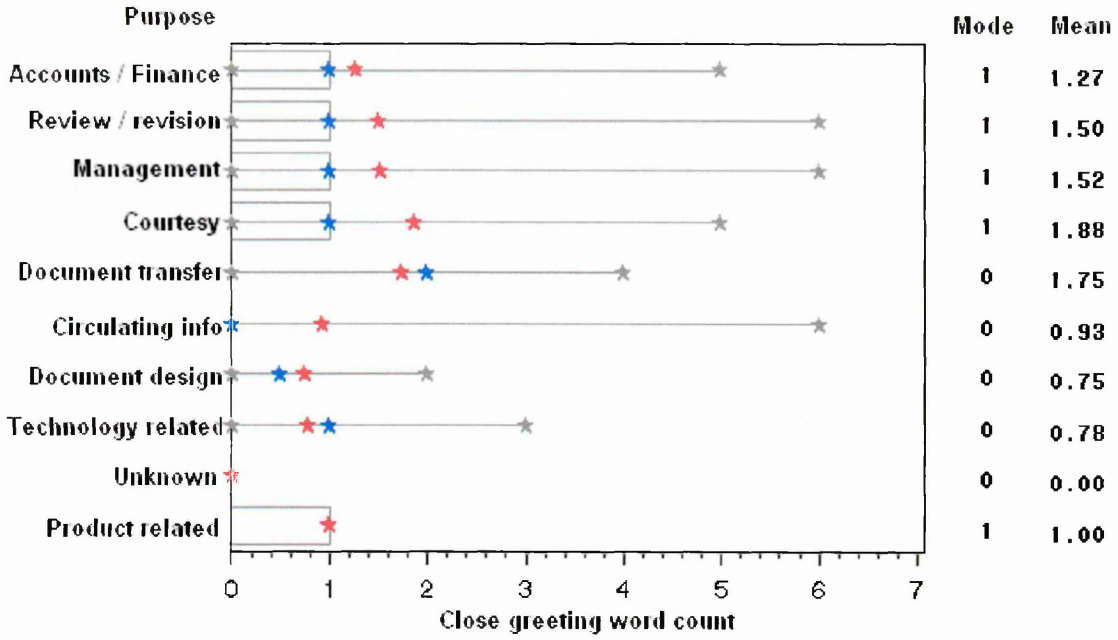

Figure 5-10: Differences by purpose for close greeting word count

Legend: min ★— max —★ median ★ mean ★ mode □

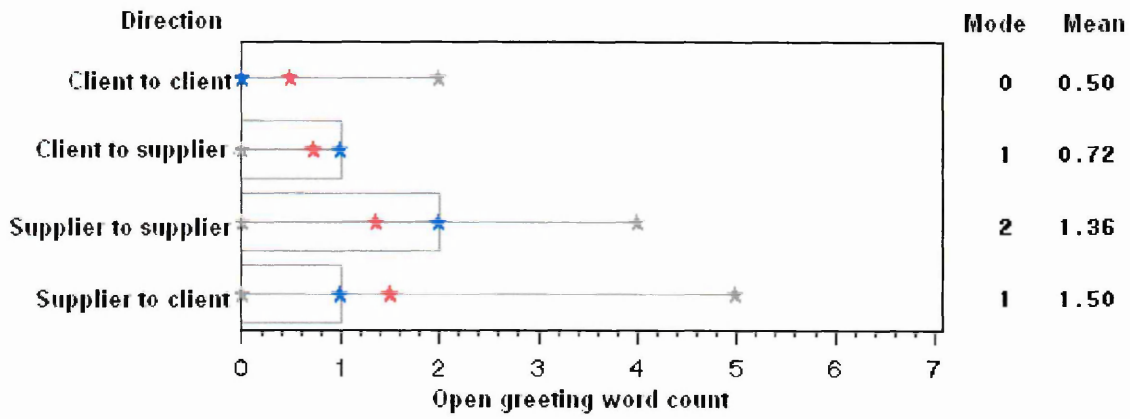

Figure 5-11: Differences by direction for open greeting word count

Legend: min ★— max —★ median ★ mean ★ mode □

Open and close greetings also differed with relational direction of an email (see Figure 5-11 and Figure 5-12). Interpreting the three modes for open greeting length as three levels of formality, we see from Figure 5-11 that the client in house emails are least formal, the most frequent behaviour being not to include a greeting. The most formal open greeting style is practised by the supplier team members in house, possibly accounted for by the inclusion of less familiar remote contractors working on the supplier side (see Figure 3-2). Although modes for the client to supplier and supplier to client open greetings are equivalent, the higher range and mean for

the supplier to client emails suggests that these emails tend to be more formal than the reverse emails, a trend strongly portrayed in the close greetings (discussed below).

The client interviewee commented on open greeting results, confirming an expectation of increased formality in emails going outside the organization as follows:

> *Colleagues internally always write differently to when addressing external individuals (client: appendix O, line 18).*

The interviewees further justified variations in open greeting lengths by organizational norms. The supplier interviewee explained as follows:

> *Banksys [the client] always use the first name; at Namahn we usually say more than just the first name and we use the same style with subcontractors e.g. Dag Peter, Hi John (supplier: appendix N, line 25).*

The client interviewee referred to possible norms in the organizational cultures, which may be reflected in communication behaviours:

> *It's surprising that the client to client vary from the supplier to supplier communications. Maybe this is because Namahn have a user centred philosophy..... We are perhaps more technically oriented, more to the point, more short. (client: appendix O, line 20).*

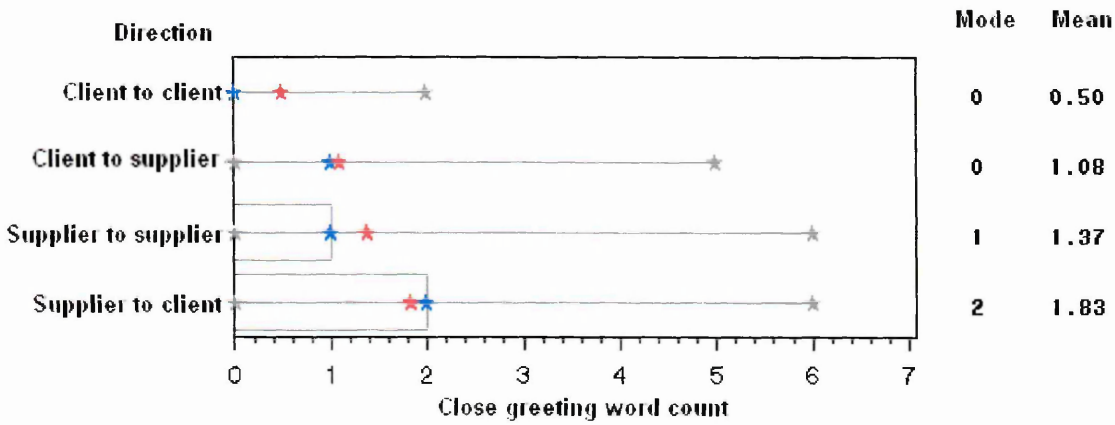

Figure 5-12: Differences by direction for close greeting word count

Legend: min ★— max —★ median ★ mean ★ mode □

For close greetings (see Figure 5-12), the most frequent practice on the client side both with emails sent in house and externally was omission of a close greeting, showing relative informality compared to the supplier side. Supplier close greetings to the client are longer and more formal than in the reverse emails, reflecting the power hierarchy in the social relationship between the two parties. This interpretation is reinforced by client interview response to question 9.

> *Namahn is a supplier and we are the client. This is a business relationship; it's different to how we are as colleagues in house (client: appendix O, line 78).*

The pattern for open greeting length with phase suggests that the first and last phases were the least formal. The supplier interviewee commented that a change in greeting formality might be expected in the second phase, when the new project leader (for the document) took over:

> *That phase two is higher is normal because in the beginning the only communications were between the product designer (later peer reviewer), and the two client team members. The technical writer and project leader was new and didn't know those people. The peer reviewer had already worked with the client for about two years (supplier: appendix N, line 54).*

Indeed, checking back in email content for dates, the technical writer and project leader took over leadership of the project from the peer reviewer on 6th November 2003, a calendar date which coincides with the beginning of socialisation phase two. This explains the increase in formality during phase two, which then decreased with socialisation in phase three. The midpoint review meeting at the beginning of phase four may have renewed the formality of the client supplier relationship during the second half of Gersick's punctuated equilibrium model, accounting for the increased formality again, which finally decreased at the end of the project.

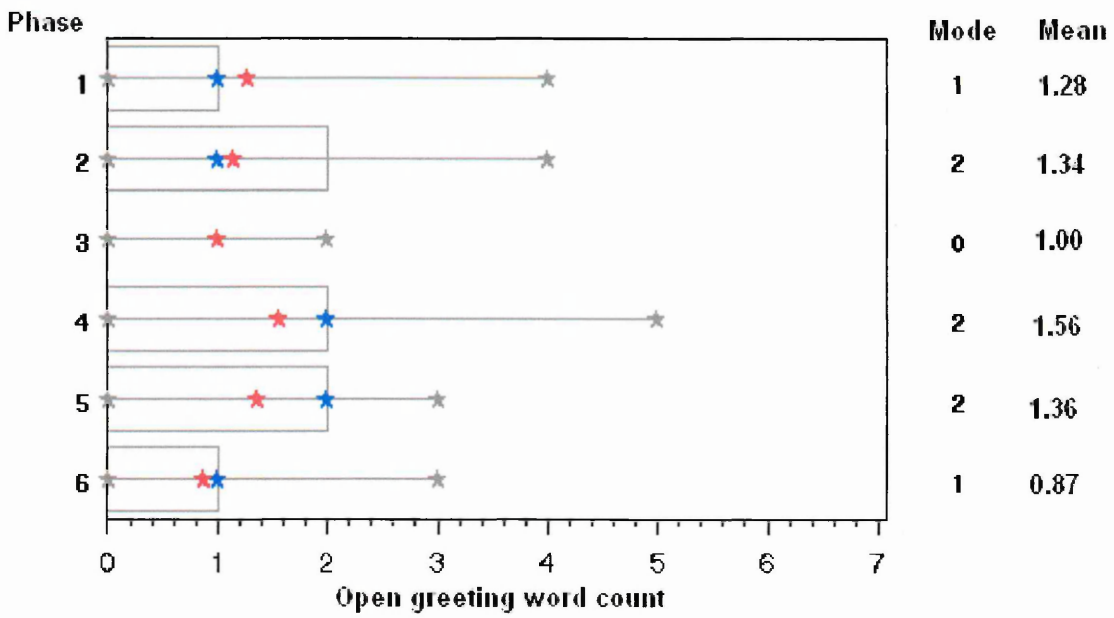

Figure 5-13: Differences by phase for open greeting word count

Legend: min max median mean mode

Open greeting length also varied according to whether the email was written in Dutch or English (Figure 5-14). Emails written in English had longer more formal open greetings, which could reflect a cultural difference or a difference in the structure of open greetings between the two languages. The latter is unlikely as the first two highest frequencies of open greeting (see Figure 3-5) are *name only* and *no greeting*, and the third and fourth are Dutch and English equivalents (*Dag* and *Hi* plus the name), which have the same length. The difference is therefore more likely to be attributable to a cultural influence on formality style as interpreted by this marker.

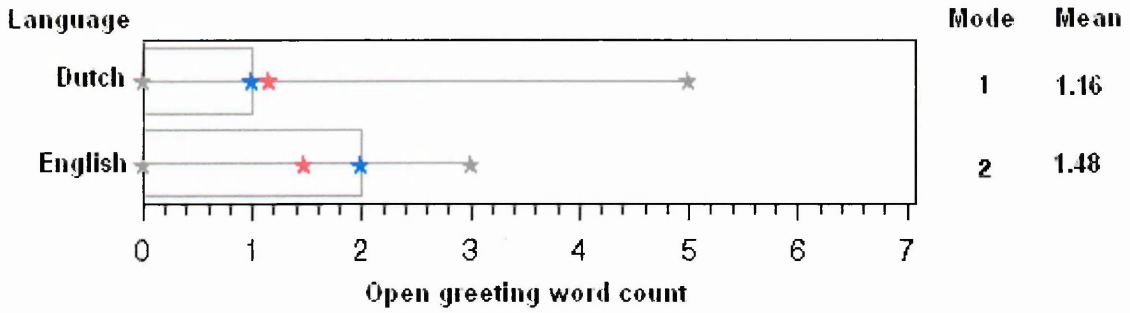

Figure 5-14: Differences by language for open greeting word count

Legend: min ★ max ★ median ★ mean ★ mode

5.3.3 Involvement interpreted from first person singular pronouns

First person singular pronoun frequency as a percentage of the total body text word count per email was used as a marker of a person's sense of involvement both in the task and the relationship with the *Receiver*. Frequency of the involvement marker was higher than the solidarity marker (involvement total count 270, solidarity total count 121 in 218 emails).

This marker showed significant differences by *Sender*, *Audience size, Direction* and *Purpose*. Thus writers varied in their sense of involvement individually, and depending on the task and context of the communication.

Of particular note in Figure 5-15, is the difference between the Supplier copywriter and the other team members. This individual was the only American working on an otherwise entirely Belgian team, (see appendix J), so the difference may reflect a cultural influence.

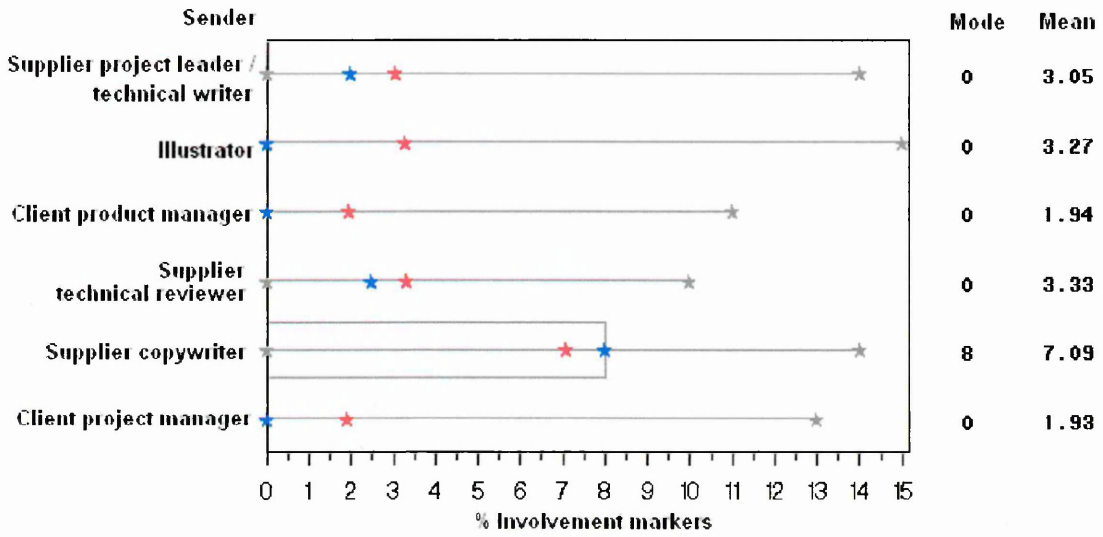

Figure 5-15: Differences by sender for involvement

Legend: min ★ max ★ median ★ mean ★ mode

Figure 5-16 shows that the range, means and medians for involvement clearly decrease with increasing audience size. Thus writers of emails on this project were more involved in private interpersonal communications than non-private interpersonal communications. This reinforces the involvement marker as being representative of socio-emotional involvement in addition to task involvement.

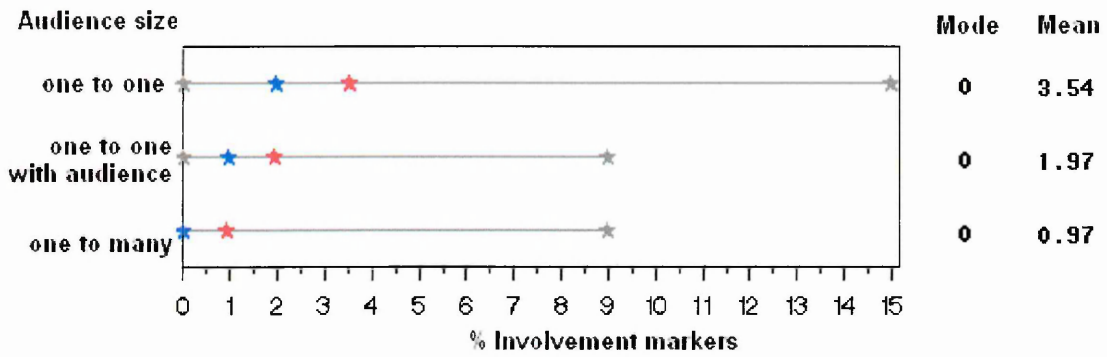

Figure 5-16: Differences by audience size for involvement

Legend: min ★ max ★ median ★ mean ★ mode

Figure 5-17 shows the differences in involvement by *Purpose*. *Courtesy* has the highest involvement. Emails written solely for courtesy purposes fulfil a socio-emotional function, to build relationships and maintain the group, and are therefore expected to have high involvement. Other purposes with relatively high means for involvement were *Accounts/finance*, *Review/revision* and *Management*. The supplier interviewee commented on preliminary results as follows:

> *Yes this seems logical to me because circulation of content may have no message and for courtesy [relatively high involvement]– yes this seems normal to me. I can't say anything about accounts, as these weren't sent by me. For document design, we were less involved. There were a few mails about whether to use their template or ours. This pattern is to be expected. Management is a little higher (supplier: appendix N, line 70).*

The client interviewee asked for an explanation of *Courtesy* mails and I responded that they were for politeness, for example, confirmation of receipt of a document.

> *Confirming receipt of documents also plays a role in management, as it advises the recipient that further activities are on schedule. This is important to aid planning. This result is a bit strange, although I suppose content circulation and document transfer are just a passing through of information, rather than a communication. There may be no content to these emails (client: appendix O, line 57).*

This last comment reflects the division of purpose types highlighted by the extremes in word counts per email. Word count was highest for tasks for which the communication content represented the task, such as *Document design* and *Review*, and lowest for *Document transfer*, where the email simply acted as a transfer agent for an attached document.

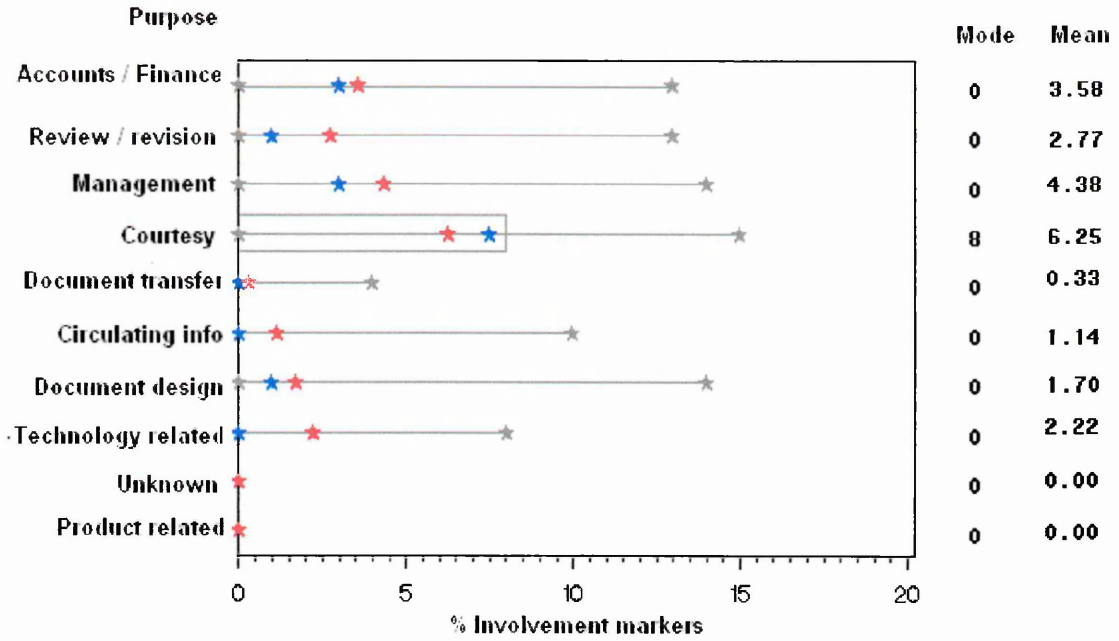

Figure 5-17: Differences by purpose for involvement

Legend: min max median mean mode

In the client supplier relationship, the client delegates work to the supplier, who completes that work, with some guidance from the client. We would expect activity and consequently involvement in the tasks to be higher on the supplier side than the client side, and this is the profile we see in Figure 5-18; emails written by authors on the supplier side have higher means, medians and ranges for involvement.

We might also expect in house emails to show more involvement than those sent externally, as we would expect a closer relationship between colleagues in an organization than across organizations. This appears to be the case in the supplier scenario, but not the case in the client scenario. A possible explanation for this is that existing relationships between client team members preclude the need for explicit representation of involvement as a pro-social communication behaviour; the involvement is assumed. In the supplier case, most of the team members were working remotely from each other, so that explicit representation of involvement was needed in the email communications. Alternatively, task and/or socio-emotional involvement between client team members may be lower than in the supplier scenario due differences in individual or organizational cultures (see interviewee comments on Figure 5-11 in section 5.3.2).

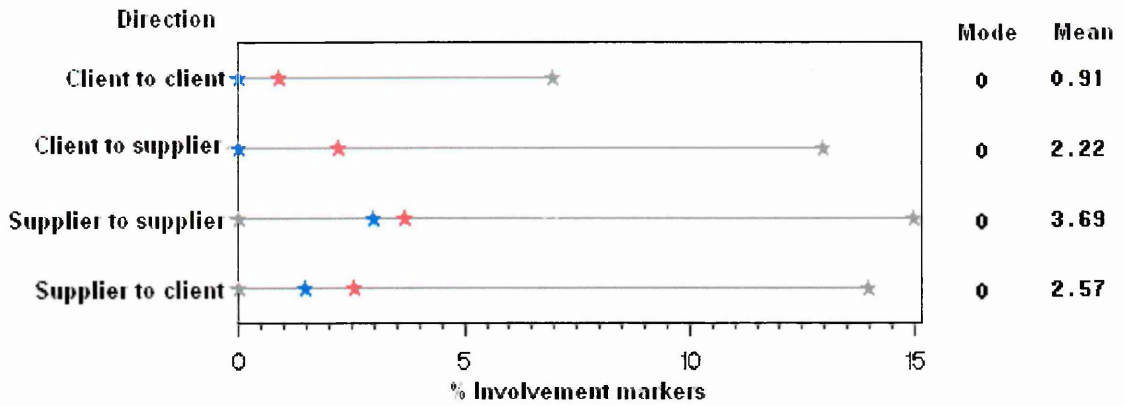

Figure 5-18: Differences by direction for involvement

Legend: min ★— max —★ median ★ mean ★ mode □

5.3.4 Solidarity and group cohesion

First person plural pronouns are interpreted as markers of solidarity or group cohesion and only varied with *Audience size*, *Purpose* and *Phase* of the project. Frequency of these markers was less than one per email (total count 121 in 218 emails).

The mean results for solidarity by audience size are shown in Figure 5-19 and suggest that solidarity was highest for interpersonal communication with an audience, than in either one to one, or one to many emails. Two thirds of the one to many communications were written from the supplier to the client, and may therefore have shown lower solidarity due to the inhibitions imposed by the client supplier relationship, or by the fact that the email was intended for recipients external to the Sender's organization.

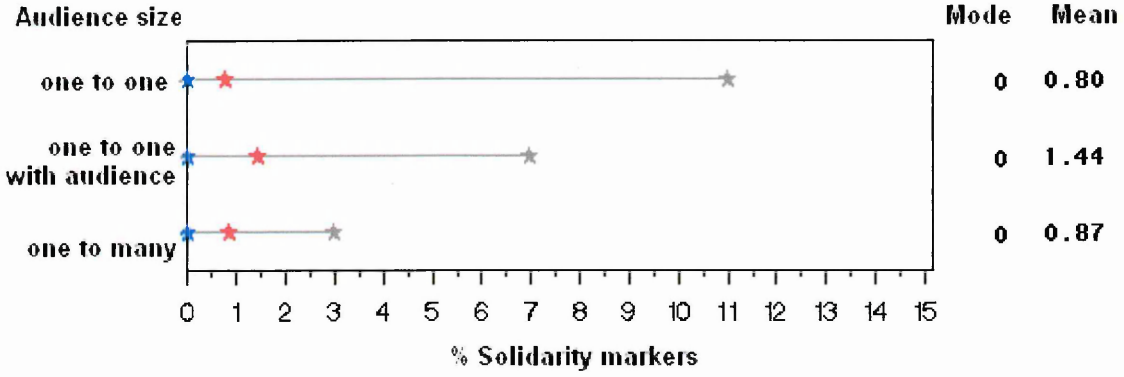

Figure 5-19: Differences by audience size for solidarity

Legend: min ★— max —★ median ★ mean ★ mode □

Secondly, involvement was highest for the one to one messages (see Figure 5-16). As use of the involvement marker and solidarity marker are mutually exclusive, we might expect high involvement to mask high solidarity. Solidarity between individuals may therefore be distorted by use of involvement markers.

Figure 5-20 shows the *Purposes* with relatively higher solidarity were *Accounts, Review, Management, Courtesy, Technology and Document design*. Interestingly the two highest means were for *Technology related* issues and *Document design*. Document design happened mostly early on (45% in phase one; see appendix II), required interactivity between the client and supplier to agree the design, and was mainly handled by the first project leader, who had already worked with the client on another project for two years (supplier interviewee: appendix N, line 57). Thus there are a number of possible reasons why this *Purpose* may show high solidarity. Further the supplier interviewee explained increased pro-social behaviour (in terms of social markers) related to the *Technology* category as follows:

> *I think it was a problem of not being able to open the illustrator's pictures at the end....if you know your subcontractor has problems, this explains the behaviour/trend (supplier appendix N, line 84).*

He thus explained pro-social behaviour for *Technology* as due to a series of emails exchanged between the supplier and a subcontractor, in which awareness of the subcontractor's application/technology related difficulties invoked sympathy and support, and led to more pro-social behaviour.

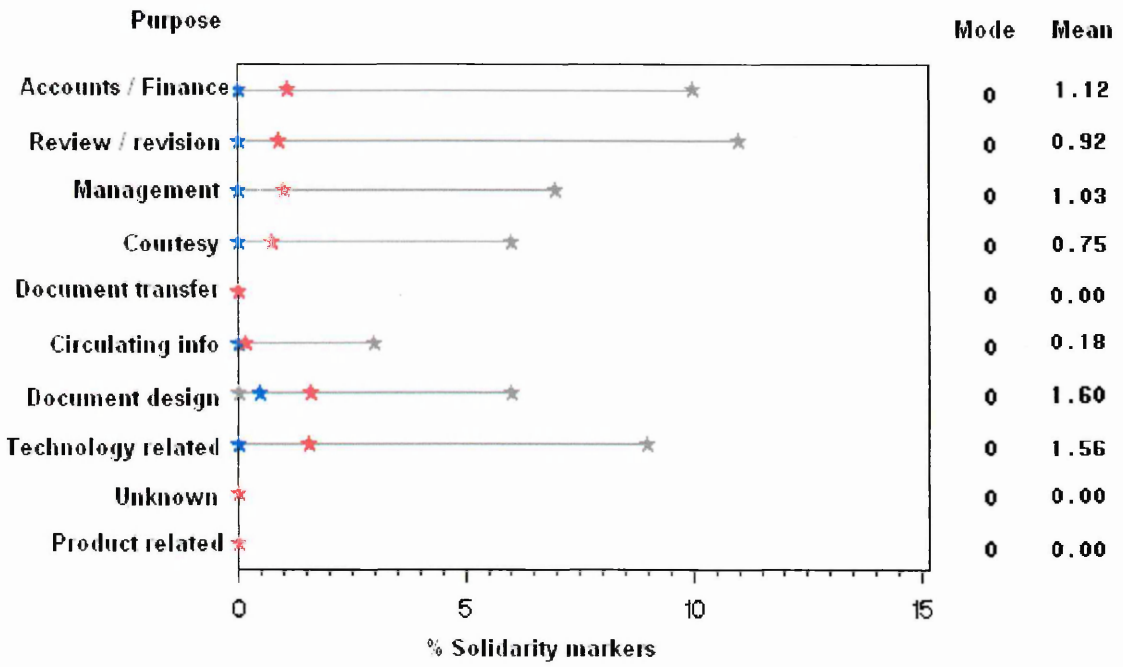

Figure 5-20: Differences by Purpose for solidarity

Legend: min — max — median ★ mean ★ mode

Figure 5-21 shows differences in solidarity by Phase. The means show a fall in the second half of the project, and then an increase at the end. This pattern is difficult to interpret. The fall in solidarity in phase two is attributable to the technical writer taking over as the new leader of the project; this individual was less familiar with the team than the previous leader. However, following the midpoint FtF review meeting at the beginning of phase four, we might expect solidarity to increase. A possible explanation for the apparent lower scores for solidarity in the second half of the project may be an increase in involvement. As already mentioned,

involvement and solidarity markers are frequently grammatically mutually exclusive, so that only one type of social behaviour can be demonstrated, either involvement or solidarity. The Kruskal-Wallis test did not show a significant difference in levels of involvement by phase, but the pattern for mean involvement to a certain extent mirrors the pattern for solidarity by phase as shown in Figure 5-22. This exclusion factor between these two markers may explain the pattern to a certain extent.

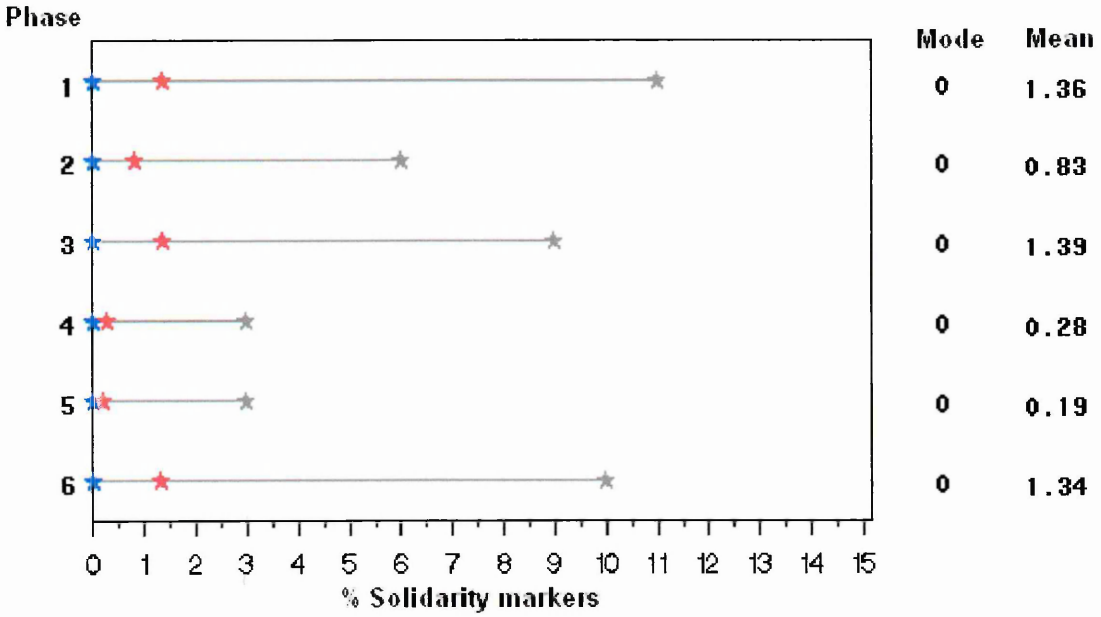

Figure 5-21: Differences by phase for solidarity

Legend: min max median mean mode

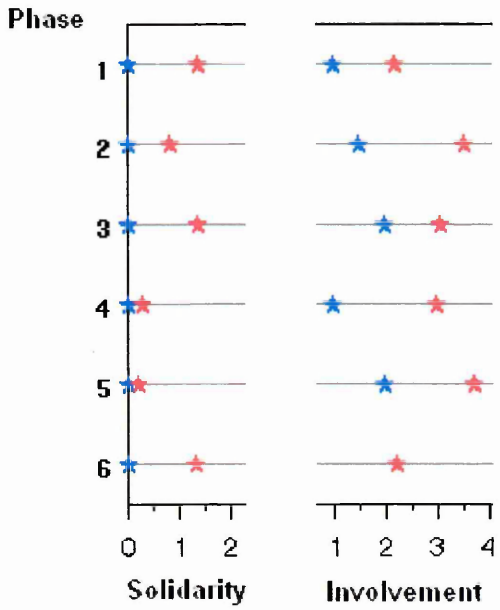

Figure 5-22: Comparison of solidarity and involvement trends by phase

Section 4 in the questionnaire (see appendix EE) probed for respondent ratings on group cohesion and sociability within the group. The ratings for all seven respondents for each question are shown in Figure 5-23. The risks of interpreting individual level data at a group level (Gallivan and Benbunan-Fich 2005; Walczuch and Watson 2001) led me to consider these respondent ratings on group cohesion in a holistic, but non-statistical way. These ratings show an overall perception of average to positive group cohesion. There are only two scores with negative valence, although it is possible that respondents may have felt inhibited to respond negatively, which might be interpreted as a criticism of other team members (Beck 1993 p109). Respondents 1 and 6 answered the first question: "did you feel that you were really a part of this team?" with the response "didn't work in a particular team" (score 0). The rest of the scores were positive, which indicates an overall tendency towards a positive interpretation of group cohesion from the main actors in the group (seven out of the eight respondents completed the questionnaire.)

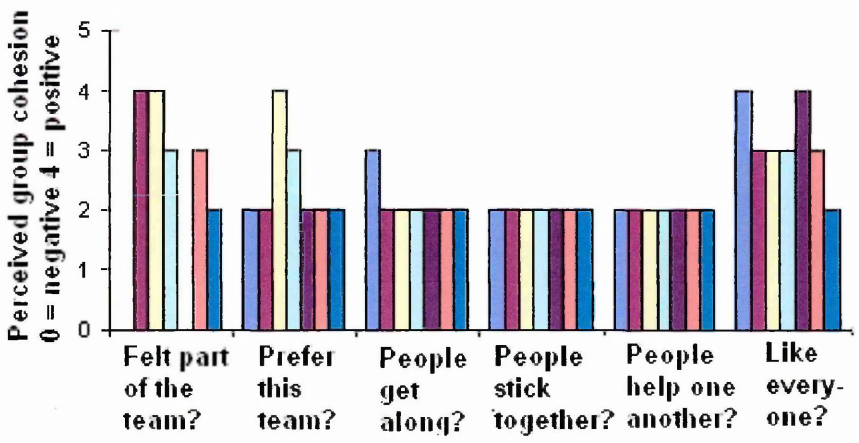

Figure 5-23: Perceived group cohesion (seven respondents, six questions)

The supplier and client interviewees were asked to comment on the interpretation that group cohesiveness was perceived as positive, and commented as follows:

> *I didn't feel like a cohesive group, though we worked well as a team. During the project, I never saw the proof reader, though I saw the clients a few times, especially at the beginning. I spent two days with the client at the beginning, at the project manager's desk. I almost didn't see the peer reviewer at all during the project, even though he was a Namahn team member. It didn't feel like a cohesive group to me because I didn't meet people face to face. I don't think it's possible to build the same relationships by email, but maybe I'm old fashioned; for me the best communication is face to face (supplier: appendix N line 101).*
>
> *Yes I would agree with this (client: appendix O, line 91).*

Thus in spite of the general positive perception of group cohesiveness portrayed in Figure 5-23, one team member on the supplier side had reservations.

5.3.5 Formality score: greeting and signature styles

The formality score was derived from quality of greetings and manual signature, and presence of automated signature. The Kruskal-Wallis test (Table 5-1) shows differences between values for the formality scores for all the independent variables except *Audience* and *Language*, and I discuss the results here.

Figure 5-24 shows that two team members wrote emails rated at the central value and the rest were rated as slightly less formal. Figure 5-25 shows that writers also adapted their level of formality according to the person they were writing to. Clients and subcontractors (illustrator and copywriter; see Figure 3-2) received more formal communications than the supplier team members.

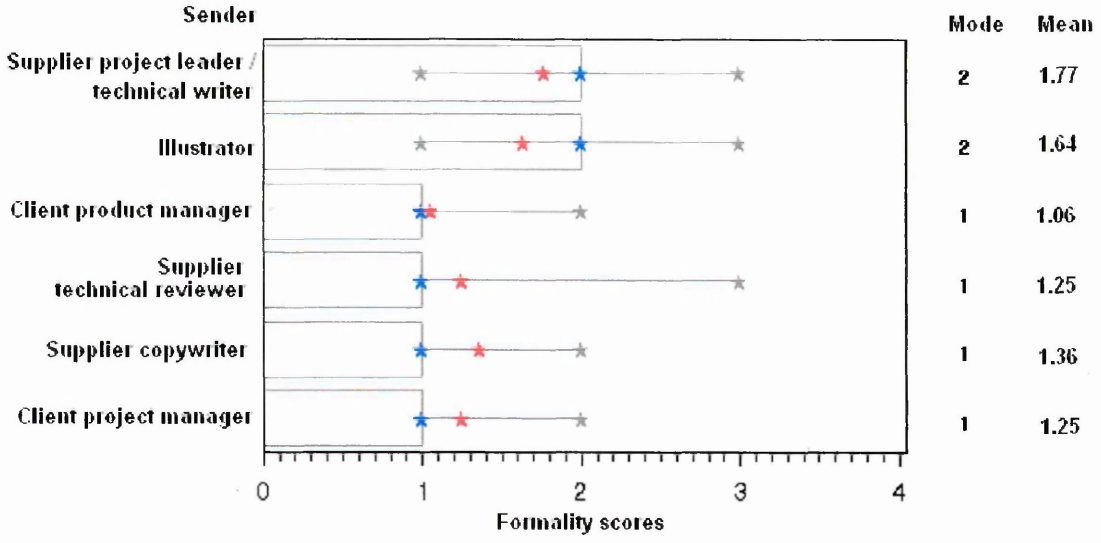

Figure 5-24: Differences by sender for formality scores

Legend: min max median mean mode

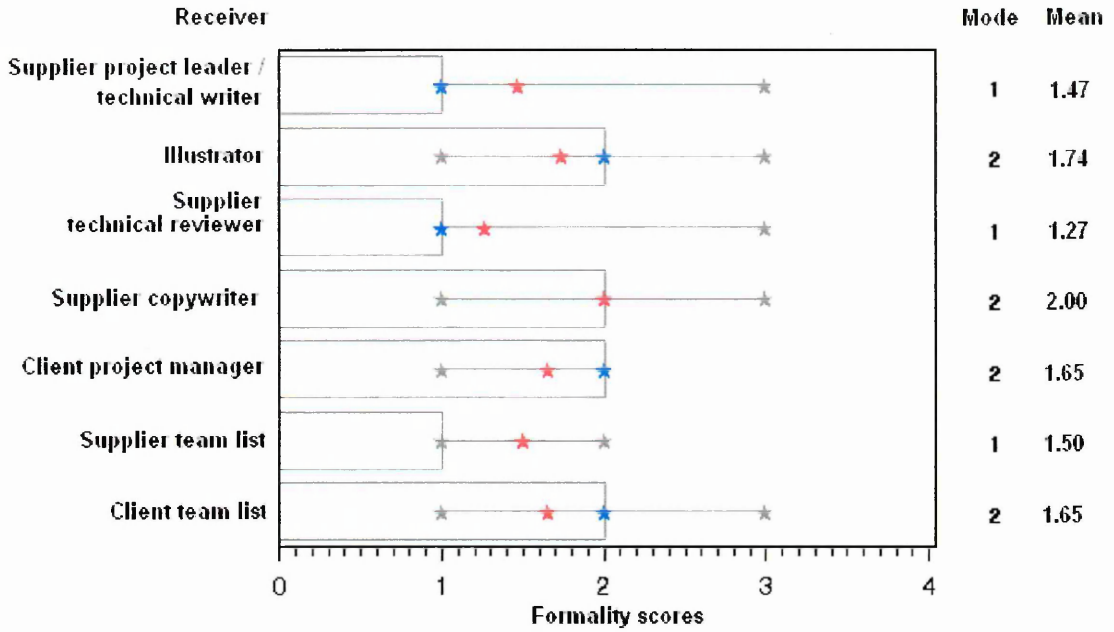

Figure 5-25: Differences by receiver for formality scores

Legend: min ★— max —★ median ★ mean ★ mode □

Communications on *Accounts*, to *Circulate information* and *Document design* scored as less formal than other purposes (see Figure 5-26). This result for *Accounts* conflicts with the client respondent's perception of early tasks such as *Accounts*, requiring longer more official greetings, at a time when people didn't know each other early in the project. However, *Document Design* emails may have been less formal as these involved the first project leader, who was already familiar with the client contacts.

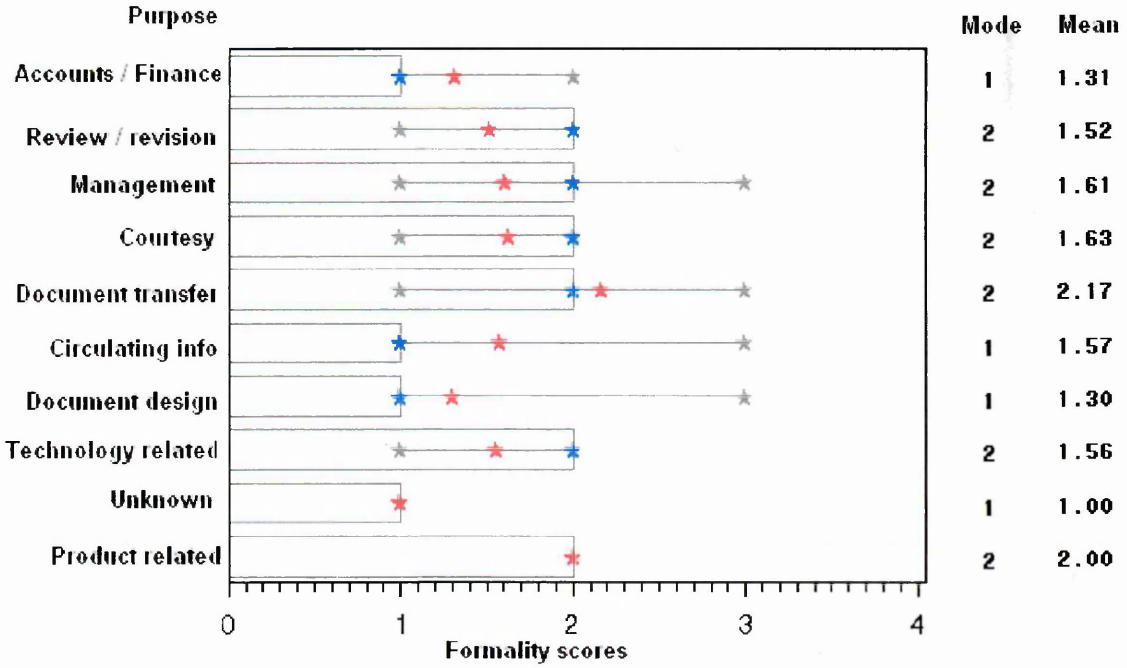

Figure 5-26: Differences by purpose for formality scores

Legend: min ★— max —★ median ★ mean ★ mode □

A decrease in formality might be expected throughout the project, although the supplier interviewee pointed out that team membership change in phase two might account for a slight increase in formality and this is shown in Figure 5-27. The reduction in formality at the end is explained by communications becoming less formal, reflecting closer relationships and group cohesion with socialisation.

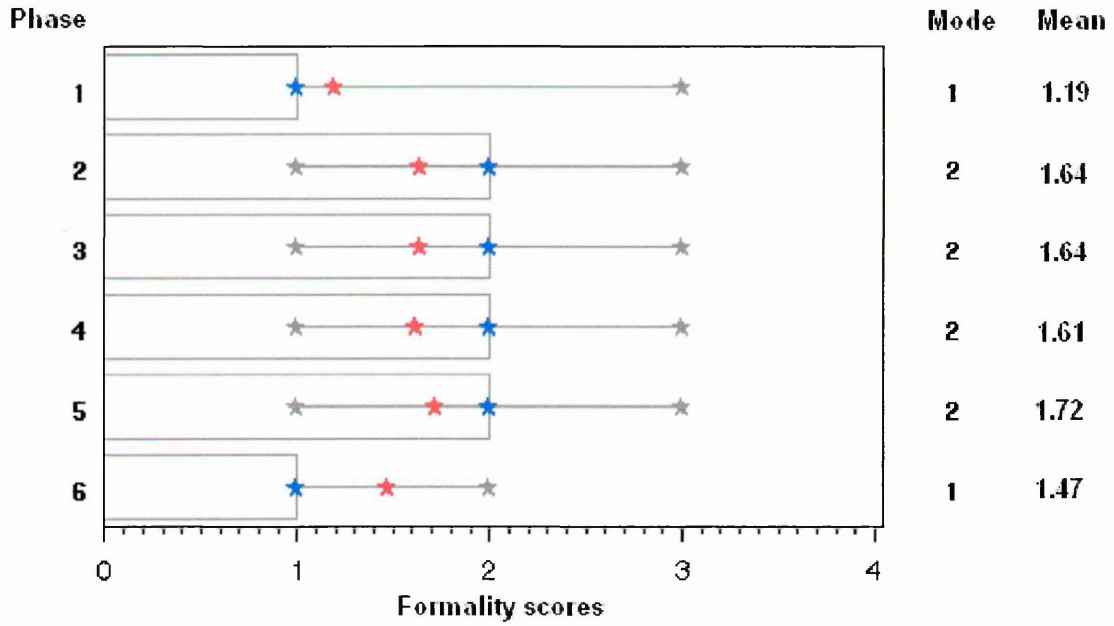

Figure 5-27: Differences by phase for formality scores

Legend: min max median mean mode

Supplier communications were more formal than client communications (Figure 5-28). This latter point is reinforced by open and close greeting results by *Direction*, and also corroborated by the interviewees' comments on the influence of organizational culture in response to greeting lengths, as reported earlier:

> *Banksys [the client] always use the first name; at Namahn we usually say more than just the first name and we use the same style with subcontractors e.g. Dag Peter, Hi John (supplier: appendix N, line 25).*
>
> *It's surprising that the client to client vary from the supplier to supplier communications. Maybe this is because Namahn have a user centred philosophy….. We are perhaps more technically oriented, more to the point, more short (client: appendix O, line 20).*

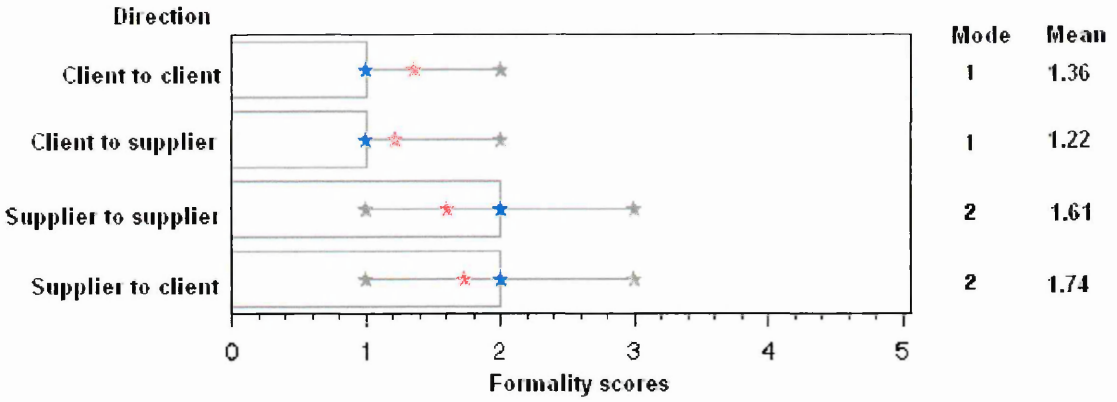

Figure 5-28: Differences by direction for formality scores

Legend: min ★— max —★ median ★ mean ★ mode □

5.3.6 Sociability interpreted from social building units

Texts which might contribute towards the social development, interpersonal relationships and maintenance of the group were interpretively coded as "social building units". The total number of social building units per email varied with *Receiver, Audience size, Purpose* and *Language.*

Figure 5-29 shows mean and median values were highest for the illustrator and supplier copywriter *Receivers*, suggesting that these two individuals may have experienced the closest relationships or been the most liked in the teams.

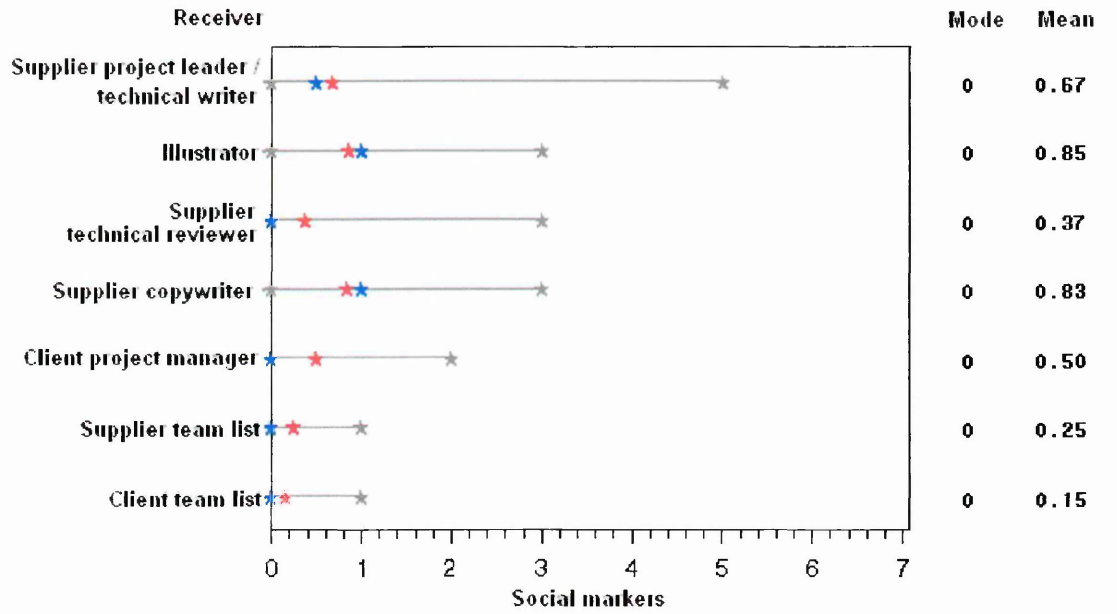

Figure 5-29: Differences by receiver for social markers

Legend: min ★— max —★ median ★ mean ★ mode □

The Kruskal-Wallis test showed that social marker frequency differed significantly between at least two categories of audience size. Ranges and means in Figure 5-30 show that frequency of

social building markers decreased with audience size, reflecting the closeness and privacy of the interpersonal communication.

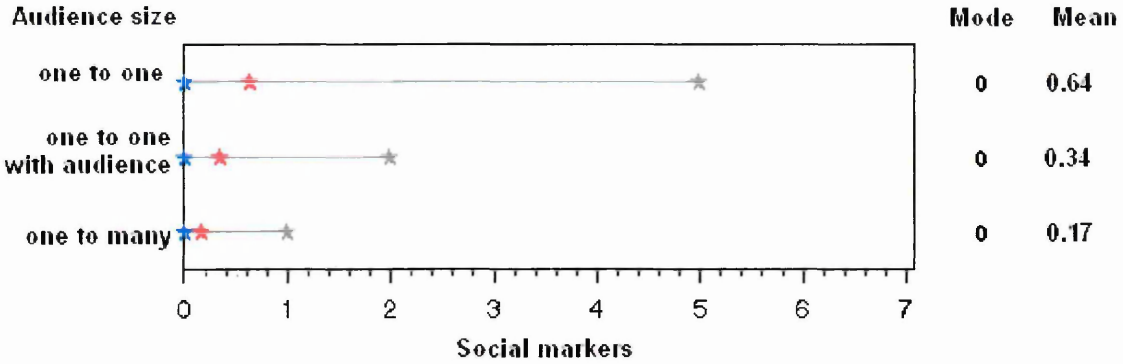

Figure 5-30: Differences by audience size for social markers

Legend: min max median mean mode

Purposes (see Figure 5-31) with relatively higher mode and median scores for frequency of social markers are *Courtesy* and *Technology related* issues. The supplier interviewee commented on preliminary results on social markers as follows:

> *For courtesy messages [to be higher] yes this seems logical. For technology yes and no.....if you know your subcontractor has problems, this explains the behaviour/trend (supplier appendix N, line 84).*

The interviewee confirms that he expects *Courtesy* mails to have a higher number of social building units, as the overall aim of the email is pro-social. He explains the pro-social communication behaviour for *Technology* was probably due to a series of emails in which his awareness of the subcontractor's application/technology related difficulties invoked his sympathy and support, and led to more pro-social behaviour. Thus the interviewee's comments also explain the high sociability score for the Illustrator subcontractor in Figure 5-29.

The client interviewee responded:

> *Yes, this is logical. I wouldn't expect things to be so sociable at the beginning when accounts and management issues are being discussed. This is a period of negotiation to see if you can do business together. I would expect this to vary with later tasks in the project, such as review discussions (client: appendix O, line 70).*

The examples the interviewees give explain the trend of earlier tasks such as *Accounts*, having lower sociability than later tasks such as *Review*.

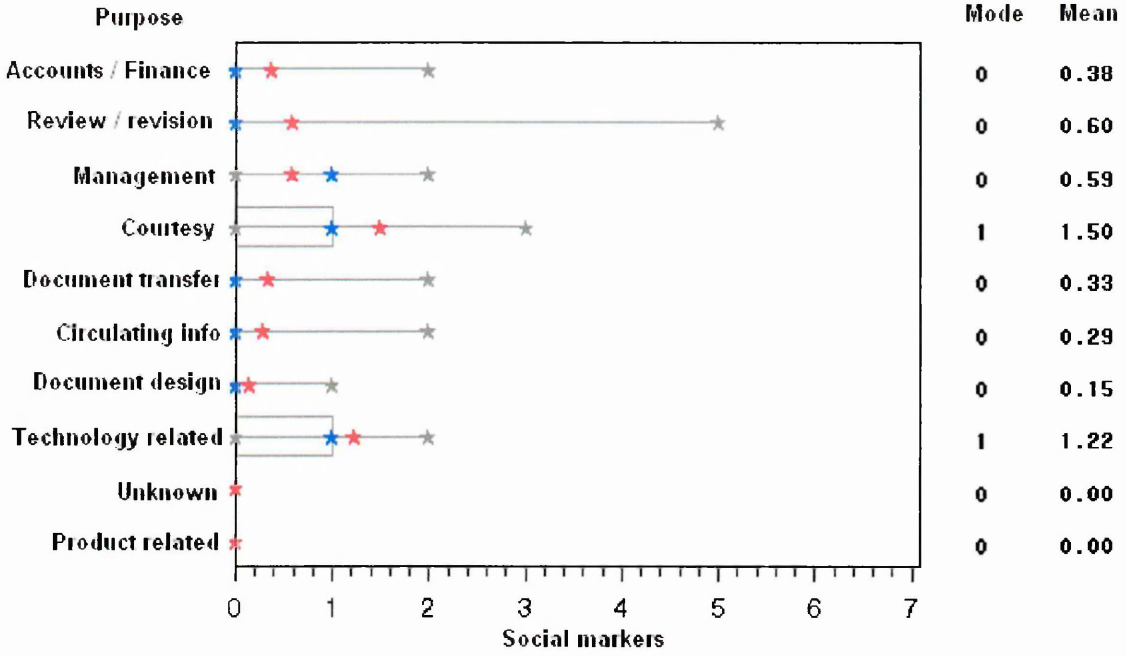

Figure 5-31: Differences by purpose for number of social building units

Legend: min ★— max —★ median ★ mean ★ mode □

Of particular note during the coding was the lack of self-disclosure of personal information. Surprisingly, sociability did not vary by *Sender*, *Direction* or *Phase* and a clue to the reasons for this may lie in the components of the marker. Different subcomponents of the social markers may represent slightly different communication behaviours, which therefore confuses the results. The social building texts were further categorized as politeness, social building, apology or humour. The frequencies are listed in Table 5-2 and total 116 units in 218 emails. Distributions of the separate types of social building units across the socialisation phases are shown in Figure 5-32.

Table 5-2: Frequencies of social markers interpretively coded into categories

Social Marker	Frequency	Examples
Politeness/courtesy	71	"Thanks," "Bedankt," (thank you) "Alstublieft" (please)
Social building	37	"good holiday" "see you soon" "good luck with the event "questions, questions....." "looking forward to working with you" "it was nice meeting you in person" "all the best in 2004!" "well tried"
Apology	6	"sorry, almost forgot"
Humour	2	"(now who would have thought that)" "I hope the size of your reference number doesn't reflect how much work you've done!"

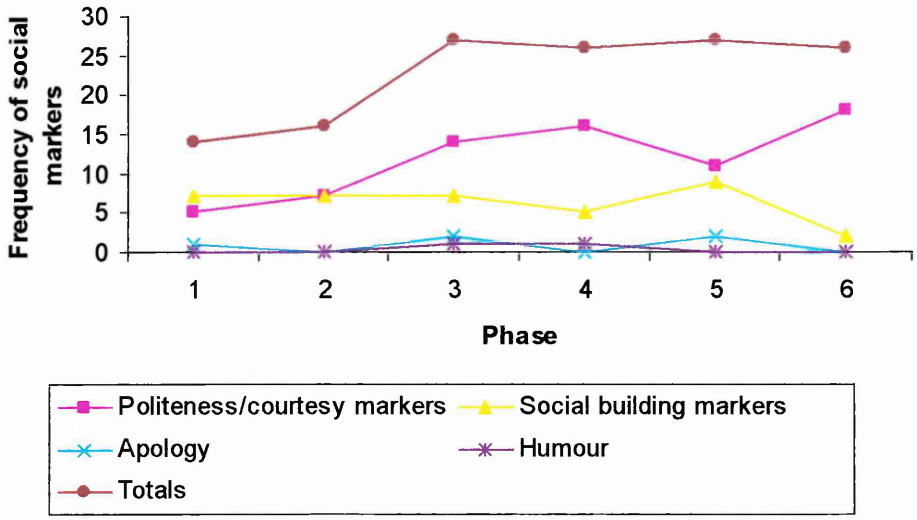

Figure 5-32: Distributions of social building units by socialisation phase (116 social building units; 218 emails)

The interviewees' comments reinforce the pattern shown in Figure 5-32:

> *I would expect sociability to increase through the project, while you are getting to know your colleagues. (supplier: appendix N, line 97).*

If the phases were time based, I would expect seasonal issues, such as Christmas to influence this….There may be more sociability early on to get the project started, and then it might remain pretty much the same (client: appendix O, line 84).

Indeed, Christmas and New Year both fell in phase five, which accounts for the increase in general social building units, and the overall sociability does seem to increase until around the middle of the project and then remain relatively stable. Of particular note, however, is the difference between the directional trends of the general social building markers, and politeness markers. Overall, general social markers decrease with socialisation phase (apart from during phase five, which included Christmas and New Year (and therefore additional expressions of seasonal wishes), whereas politeness markers increase throughout the project. This is particularly interesting, because it not only highlights the different representations of component social markers, but also the importance of understanding the situational representation of such markers. Politeness and courtesy may be erroneously interpreted as formality markers, but in this project their frequency increases with socialisation phase; i.e. with increasing familiarity individuals showed increasing courtesy. This might be explained by reduced inhibition and an increased tendency to include informal socio-emotional expressions of gratitude.

The number of social building units did not vary significantly with relational direction (client to client or client to supplier etc.) in this project. However, a preliminary analysis (prior to the interviews) had suggested that social building unit frequency did vary with relational direction and interviewees commented on this as follows:

Of course this is different. It's normal. A client is different from a colleague (Supplier: appendix N, line 92).

This is normal. We always communicate differently with colleagues in house to how we communicate with external professional contacts. Namahn is a supplier and we are the client. This is a business relationship… (Client: appendix O, line 77).

Thus, the interviewees expected two influences to affect sociability, the client supplier business relationship and belonging to separate organizations ("we" and "external professional contacts" in the above quotation).

Finally, English emails (Figure 5-33) had higher values for the mean and median frequencies of social markers than Dutch emails on this project, which may reflect a cultural difference.

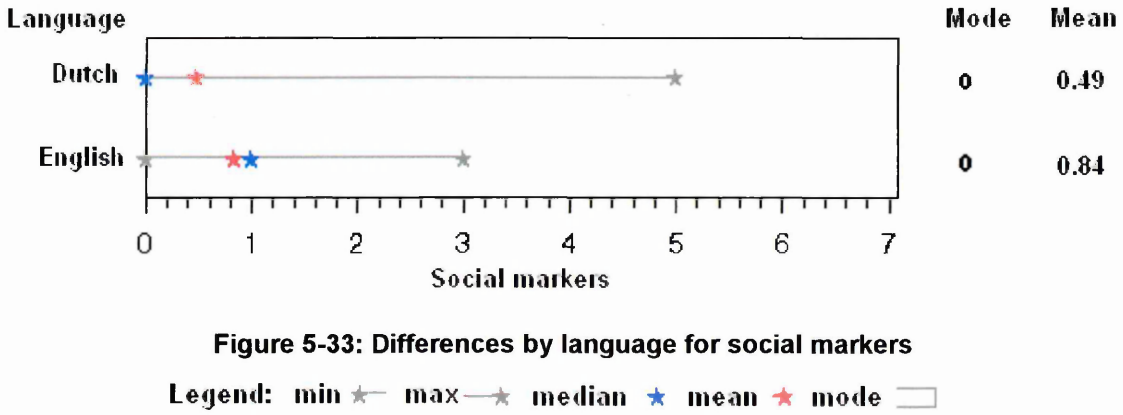

Figure 5-33: Differences by language for social markers

5.4 Discussion

5.4.1 Overview

In the previous section, I considered each of the communication behaviours which showed significant variations with the writing influences. All the writing influences studied in this project caused significant variations in at least two of the communication behaviour markers used (see Table 5-1). The most prevalent influence was *Purpose*, which affected all the communication behaviours, and represents the task dimension of the project.

Detailed analysis of the variations within categories of writing influences together with interviewee and questionnaire data helped to develop meaningful interpretations of some of the task and social dynamics represented by the communication markers. In this section, I draw together the results from different communication markers to build an overall profile of the balance between the task and social dimensions of the project. First I consider the results mapped against Nystrand's social interactive model of written communication, and then discuss the evidence gathered pertaining to group cohesion. I then present overviews of multiple communication markers, which combined reinforce some of the interpretations discussed in the previous section. Finally I comment on the indiscreteness of the social and task dimensions, and the complexity of the markers, which require holistic rather than discrete interpretations.

5.4.2 Email style and the social interactive model of writing

The aim of this research is to design an email analysis tool, which can be used to understand how team culture influences virtual team writing. Networked team members on writing projects need to communicate by email to achieve their team objective. Evaluating communication behaviour demonstrated in team's written emails serves as a proxy means of predicting social interactive adaptations for audience, purpose and context of use in the final written document. Knowing that the social dimension of team projects impacts performance, this research focuses on the communication markers which represent both task and socio-emotional components of emails.

In Nystrand's social interactive model of writing, writers anticipate readers' needs, and meaning and interpretation is a shared social reality, the meeting of writer intentions and reader interpretations. Of surprising note in this study is that only three of the communication markers representing social behaviour, word count, formality and social markers, showed category differences for the writing variables by *Receiver*. Thus elaboration, formality and pro-social communication behaviours were used to adapt email style to help create a shared understanding between the writer and reader. The team did not vary greeting style, involvement or solidarity to achieve a shared understanding with *Receivers*.

Lack of adaptation of greeting style, involvement or solidarity for *Receivers* may be explained by norms. We have seen in Figure 3-5 to Figure 3-6, that there are clear norms of behaviour in greeting style in emails on this project; social conformity may therefore inhibit individuals from

adapting their greetings extensively for an individual *Receiver*. There is also evidence to suggest that organizational and functional factors are influential on greeting style and involvement. The Kruskal-Wallis tests showed significant differences for both open and close greetings and involvement with both *Direction* and *Purpose*. *Purpose* also affected adaptations in solidarity.

Involvement may reflect one or both of the task and social dimensions in the communication. Although writers adapted involvement with the *Audience* size, they also adapted involvement with *Purpose* and *Direction*, suggesting a tendency towards task involvement rather than socio-emotional involvement; involvement did not vary by *Receiver*, suggesting a lack of representation of the relational or socio-emotional element of the marker in this project.

The fact that there were only variations between *Receivers* for elaboration, formality and sociability, rather than for all of the communication markers in this study has two implications. Firstly it suggests that the balance between social and task dimensions may lean towards the task dimension, with little socio-emotional communication. I discuss this possibility further when comparing the project with an academic project in chapter 7. Secondly it highlights the complexity of the markers and their multiple and overlapping representations, which I discuss further in section 5.4.6.

5.4.3 Group cohesion and sociability

In section 5.4.2, I discussed the restricted number of communication markers used to adapt emails in anticipation of the reader in this project. Greeting style and expressions of involvement and solidarity were not adapted for individual readers. There are also other indicators suggesting that the social dimension on this project had a relatively low profile compared with the task dimension.

Solidarity scores were relatively low, with less than one marker per email (total count 121 in 218 emails). Solidarity did not vary with *Receiver,* or *Direction* which might have indicated sub groups with higher cohesion.

Although frequencies of social building units (116 markers in 218 emails) reflect that more than half of the emails had some form of "non-task" content, there were no personal references to activities outside work, i.e. there were no examples of self-disclosure in the emails. The fact that frequency of social markers varied by *Receiver* shows a sensitivity to adapt email style to the reader and this interpretation was reinforced by the supplier interviewee's feedback. The interviewee attributed the high sociability for the *Technology Purpose* as attributable to emails he had written at a time when he felt empathy for a subcontractor, who was experiencing technical difficulties.

Questionnaire responses suggested a positive perception of group cohesion, although one team member had reservations. Attributing his feelings to the leanness of the email medium, he commented:

> *It didn't feel like a cohesive group (supplier: appendix N line 101).*

5.4.4 Other social dimension interpretations

Senders

In this section, I summarize the interpretations from the detailed analysis of communication markers by *Sender*. Figure 5-34 collates the figures already shown for open and close greeting length, involvement and the formality score. These markers showed significant differences between at least two *Senders* in the team. Although these markers vary by *Sender*, showing some characterisation of email style by the individual, we also need to take into account the other potential influences on the markers, which may cause differences in style. For example, we know that open greeting in this data may also vary with *Direction*, *Purpose, Phase* and *Language*. Close greeting only varies additionally with *Direction* and *Purpose*. The differences between open and close greeting for a particular sender may therefore also be partly attributable to other influences. However, we can conclude that expressions of formality, involvement and greeting behaviour are partly individualistic in this data.

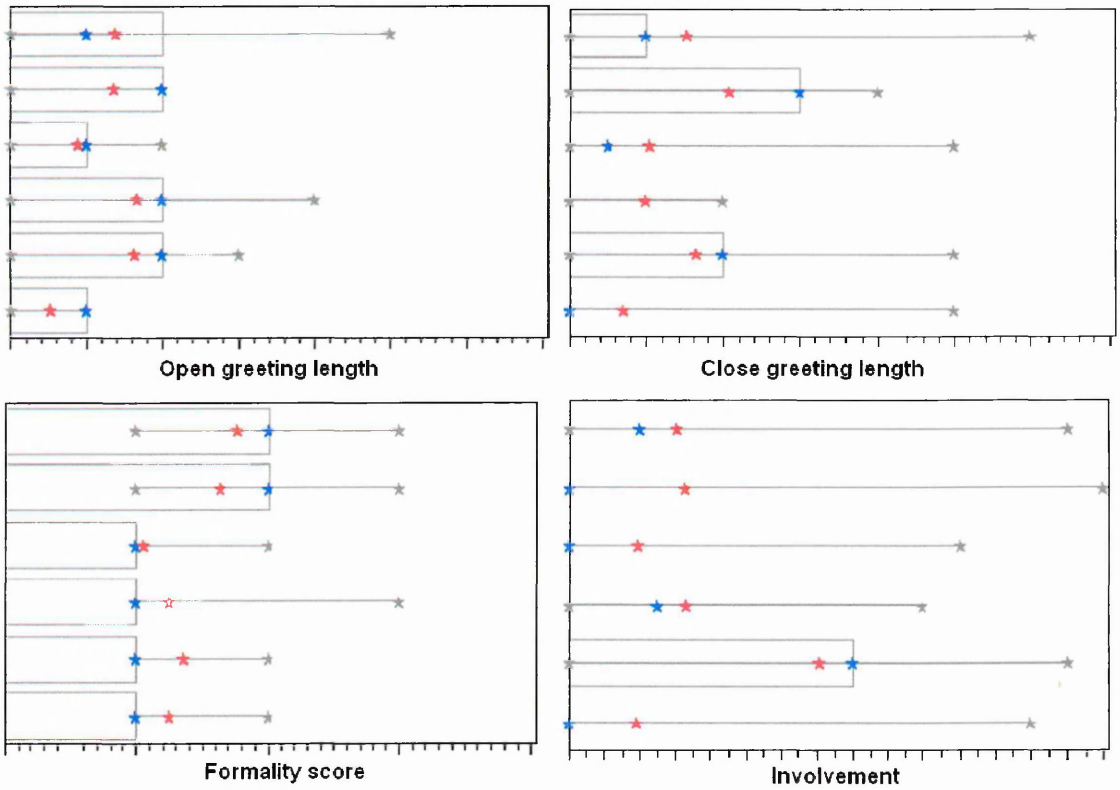

Figure 5-34: Communication behaviour by sender

Audience size

The markers provide clear information about communication behaviour trends with *Audience* size (Figure 5-35). As it became more important to write to a larger number of team members, more effort and value was attributed to emails by way of elaboration. Involvement and sociability declined with increasing audience size, reflecting the closeness and privacy of interpersonal

communication. One to many communications showing relatively lower solidarity was explained by these emails being largely to the client team, therefore crossing organizational and client-supplier boundaries.

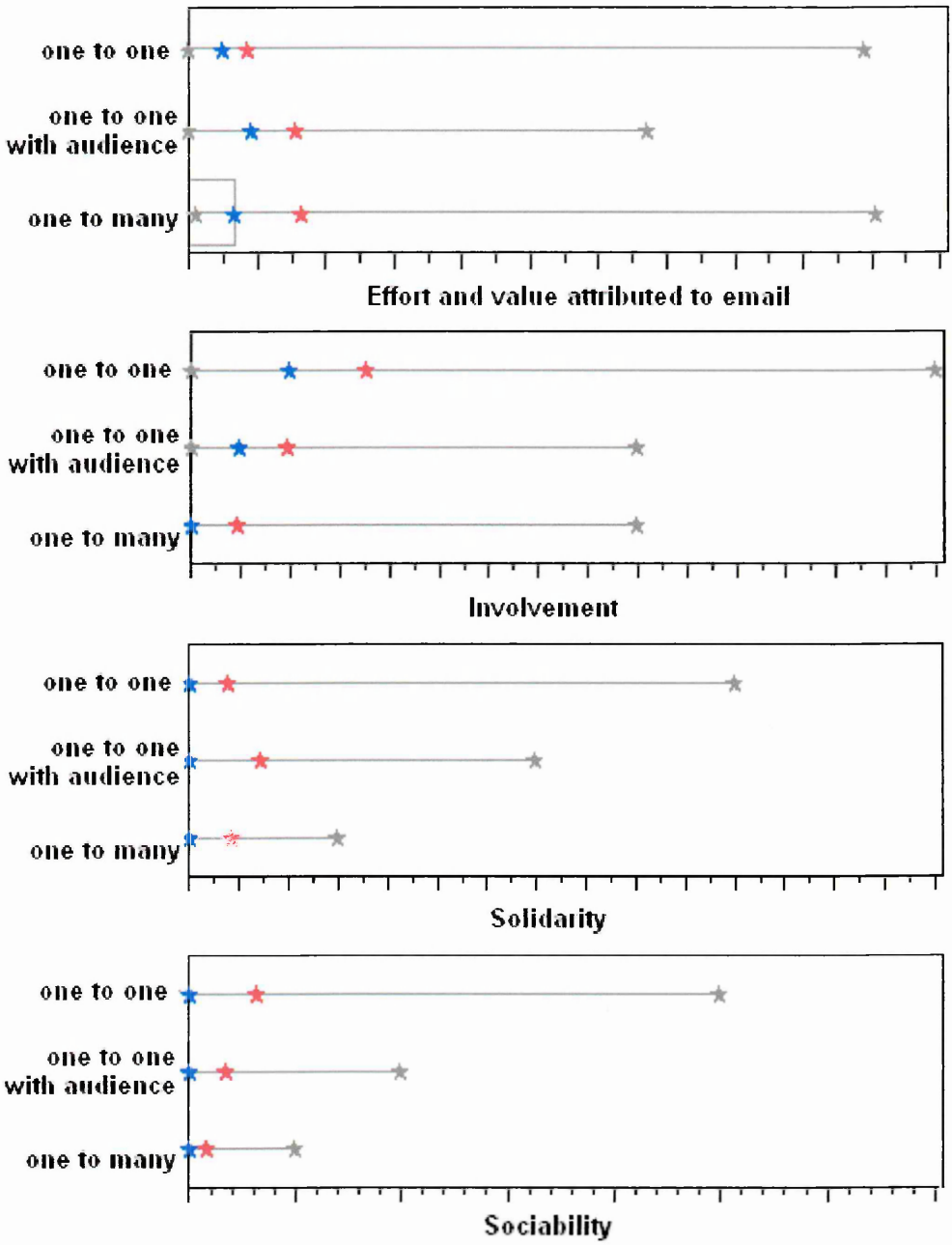

Figure 5-35: Communication behaviour by audience size

Direction

Detailed analysis of the trends by *Direction* reflects the organizational and relational differences in the communications. Figure 5-36 provides an overview of the trends for the communication markers which showed significant differences between at least two *Directions*.

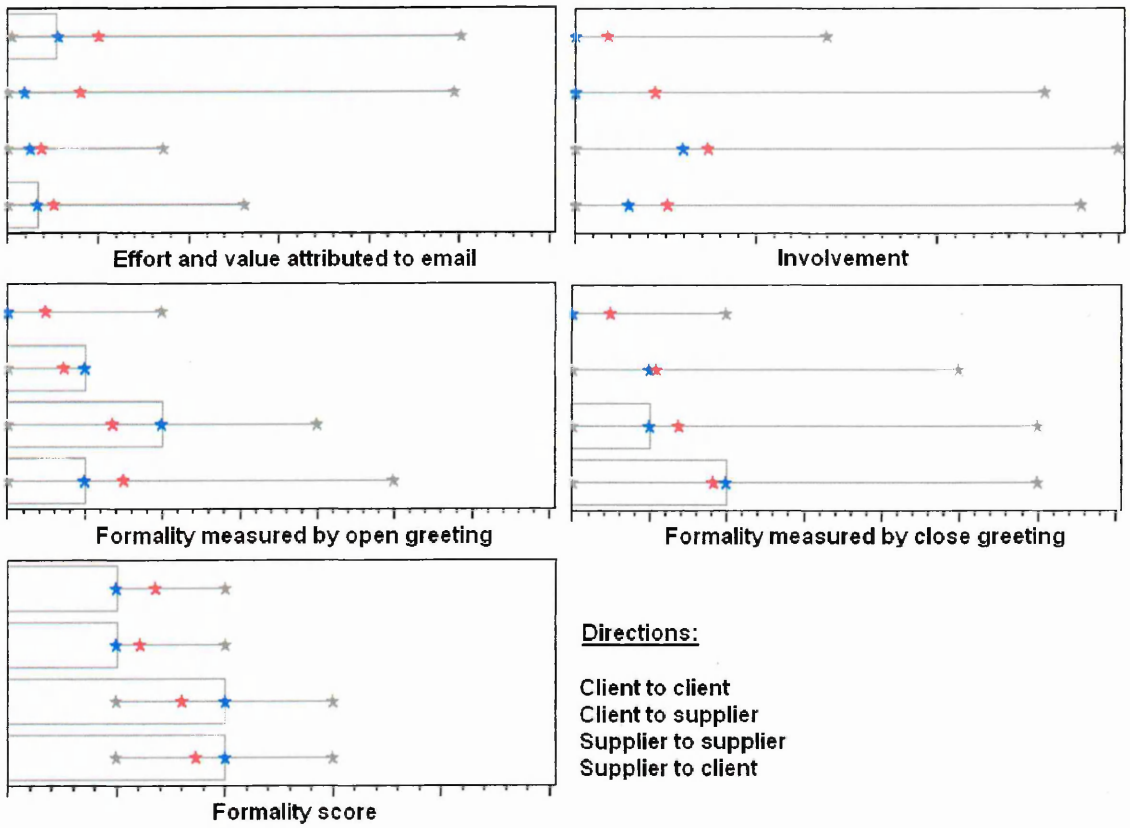

Figure 5-36: Communication behaviour by direction

In terms of elaboration, the supplier made more effort and attributed more value for communications to the client, than for communications in house, and the client made less effort and attributed less value to the communications for the supplier than for communications in house. Profiles for elaboration by relational direction thus reflected the power hierarchy of the business relationship, and we see that the equalization effects predicted by Sproull and Kiesler's (1986) lack of social context cues hypothesis are not pervasive across email communication contexts.

The client respondent summarized this as follows:

> *Namahn is a supplier and we are the client. This is a business relationship; it's different to how we are as colleagues in house (client: appendix O, line 78).*

Organizational differences were also shown in open greeting formality by *Sender*, with higher formality in supplier open greetings than client open greetings (shown in Figure 5-7). There was also a relational difference in the tendency for higher formality in open and close greetings for the supplier to client emails than for the reverse emails (Figure 5-36). Higher open greeting formality in the supplier context was explained by the interviewees as an organizational difference (see appendix N, line 25 and appendix O line 23).

Involvement was higher on the supplier side than the client side, describing the supplier-client business relationship, which defines the supplier as the most active on the project. In house (as

opposed to external) supplier interpersonal communications where relationships are likely to be more familiar also showed more involvement and less formal close greetings.

Thus the contextual relationship between writer and reader and organizational norms clearly influenced communication behaviour and email style in the team. This reflection of contextual relationships in communication behaviour corroborates Sherblom's (1988) early work, in which he showed differences in email style by level in the organizational hierarchy. It also demonstrates adaptation to the reader, albeit as defined by an organizational hierarchy, rather than the individual *Receiver*. Thus in this data, although writers appeared not to adapt greetings to *Receivers*, greetings did vary according to the writer's organizational context, and the organizational context of the *Receiver*, suggesting the team communication behaviour was socially interactive, albeit not at an interpersonal level, rather at the organizational level.

Purpose

Interestingly all the communication markers showed significant differences by purpose (see Figure 5-37). The data does not show clear trends consistently across the communication markers, although interviewees contributed some information which may explain socio-emotional components.

For example, the supplier interviewee described difficulties which a subcontractor had experienced, which had invoked his sympathy. Emails written on this topic were categorized as *Technology* related, and the supplier's concern and empathy with the subcontractor explains the high sociability and solidarity for the *Technology* category (see Figure 5-37).

We also see that *Courtesy*, a purpose which targets good relations and group maintenance, scores high on sociability and involvement (see Figure 5-37). The supplier interviewee described the *Review* purpose as the "moment of truth" which invoked anxiety, and might lead to more formality. The overview of markers for *Review* shows relatively high effort attributed to these communications, and relatively high formality in open and close greetings; thus three communication markers conform to expectations based on the interviewee's comments.

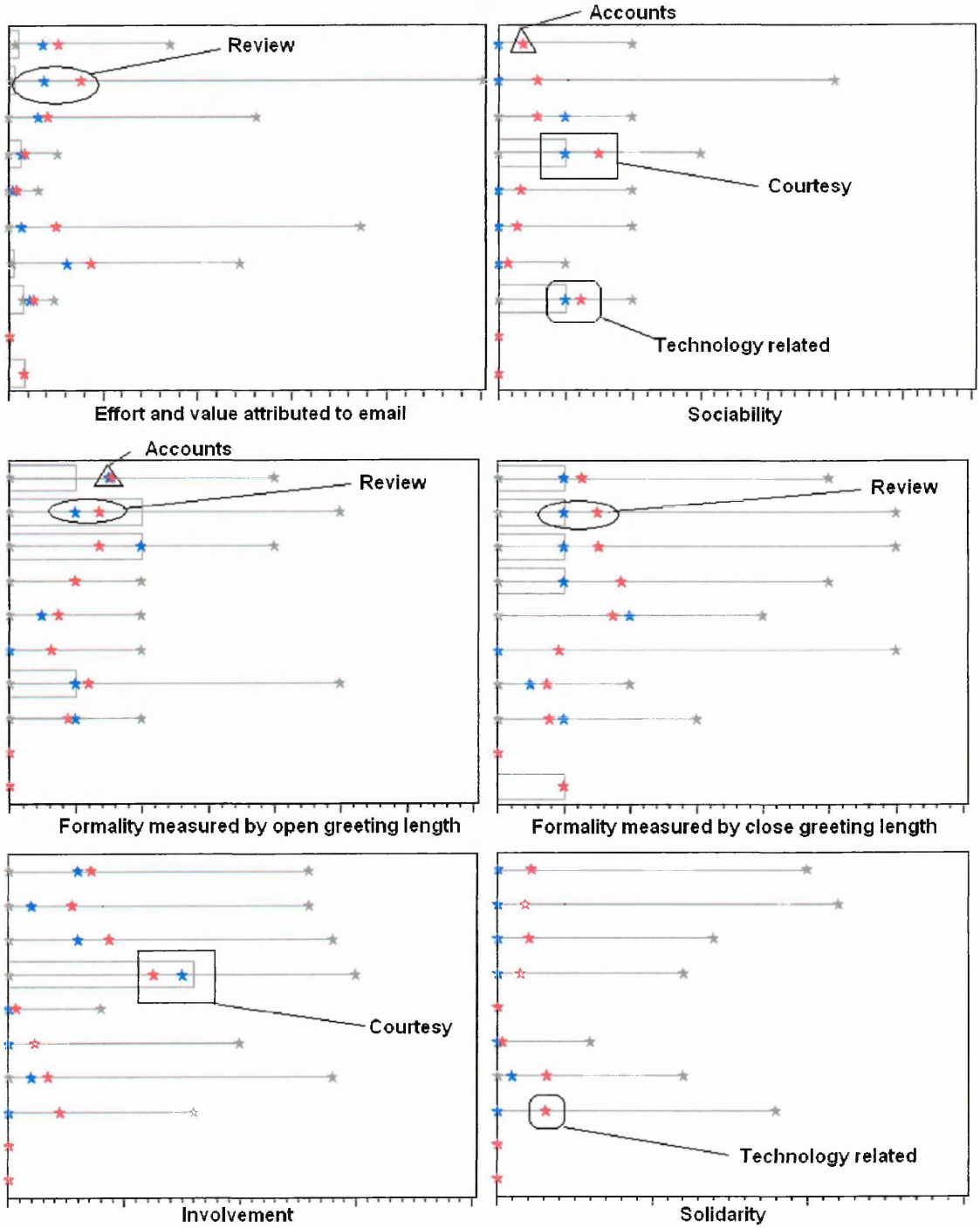

Figure 5-37: Communication behaviour by purpose

Phase

Based on social information processing theory (Walther 1995; see section 2.3.4), I expected involvement and sociability to increase with socialisation, but surprisingly, these markers showed no significant variation by socialisation phase. The dual representation of involvement, representing both task and social involvement may explain this, for example if the marker predominantly demonstrated task involvement and this component remained stable throughout the project.

Sociability not varying with phase may be explained by the varying representations of the component parts of the markers used for sociability. Overall, general social building units, such as "See you next week" decrease with socialisation phase. Politeness markers, on the other hand, increased throughout the project (shown earlier in Figure 5-32). The different dynamics in these subcomponents of the markers thus distorted the representation of sociability and masked the effect by *Phase*.

Markers which did vary with *Phase* were elaboration, solidarity, formality measured by open greeting length and the formality score (see Figure 5-38).

Effort and value attributed to the email and solidarity show peaks at the beginning, middle and end of the project. These peaks coincide with calendar dates for initial, midpoint and end FtF meetings held on the project, so that they demonstrate early team enthusiasm, renewed enthusiasm at the midpoint and towards the end of the project. Interestingly, this profile conforms to Gersick's punctuated equilibrium model which I discussed in section 2.3.2 under "Development phases". Open greeting and the formality score both indicate an increase in formality when the new project leader and technical writer took over the project in phase two, and a fall in formality at the end of the project, due to increased familiarity with team members. It seems strange that formality measured by open greeting increased in phase four, after a fall in phase three, and that solidarity also decreased in phase four. The midpoint review meeting was held at the beginning of phase four and may be responsible for this transition point and the changes interpretable from the word count, formality and solidarity markers.

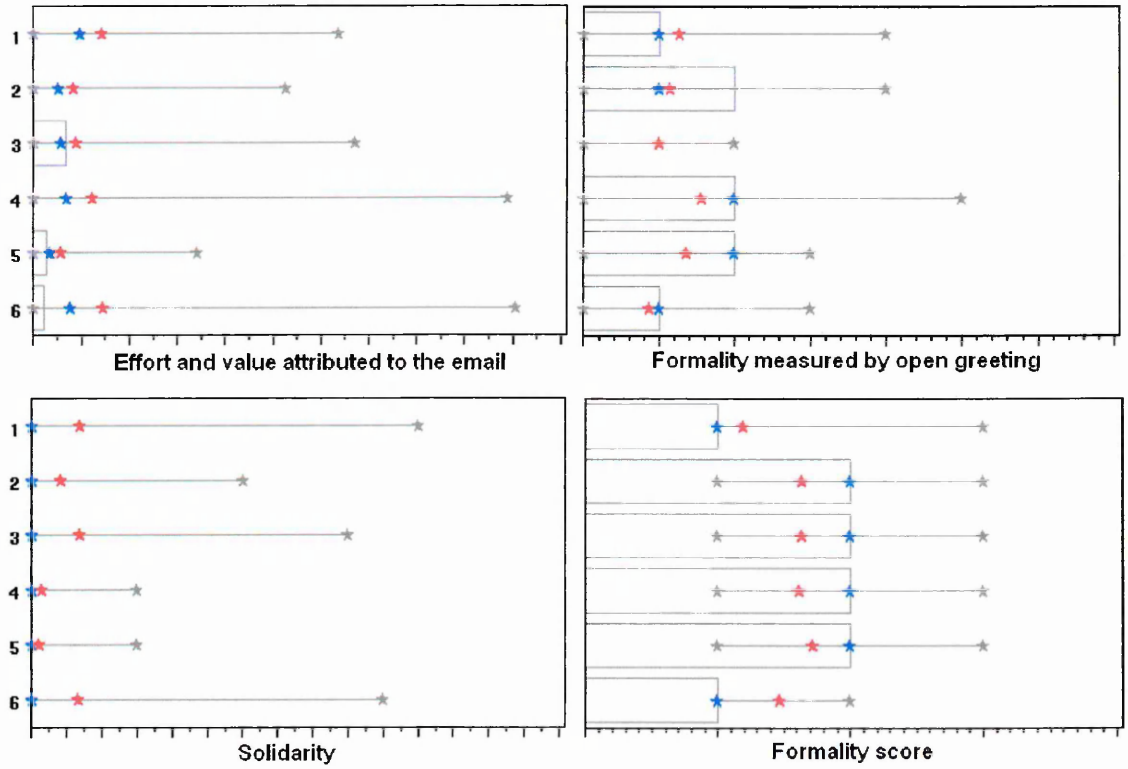

Figure 5-38: Communication behaviour by phase

5.4.5 Complex interactions between task and social dimensions

Social and task dimensions were difficult to differentiate between due to interacting factors on projects. For example, timing of *Purposes* in a project is reflected in the style of email communication. An example of this is early account-related emails showing low sociability and high formality (see Figure 5-37). Writers were more formal in communications on tasks completed early on e.g. *Accounts*, when they didn't know each other so well. Writers were also more formal in emails for *Purposes* over which they felt anxious, e.g. *Review*.

The client interviewee pointed out that even an email categorized as purely for *Courtesy* contributed towards the functional or task dimension:

> *Confirming receipt of documents also plays a role in management, as it advises the recipient that further activities are on schedule. This is important to aid planning (appendix O, line 57).*

In this example, a purpose which we might expect to be solely targeted at building social relationships and group maintenance, i.e. *Courtesy*, also serves a task-oriented objective in communicating information relevant to scheduling. Thus variations by *Purpose,* which we might expect to represent solely the task dimension, are interpreted by team participants intertwined with other writing influences, such as timing, and socio-emotional elements, such as affect.

This complexity of the interactions between the task and social dimensions of the project makes discrete interpretations from either dimension very difficult. The interpretations help towards a profile of the balance of task and social dimensions, however, which requires a holistic view of all the markers used in the study.

5.4.6 Complex marker representations

A second reason for interpreting the balance of social and task dimensions from a holistic view of the communication markers is the dual representation of some of the markers. Effort and value attributed to an email, and involvement in an email may be attributable to the task and/or social dimension, i.e. to the *Purpose* and/or relationship with the *Receiver*. In this project elaboration varied both with the *Receiver* and the *Purpose*, suggesting that effort and value were attributed both to the task and to the relationship, i.e. to both the task and social dimensions. On the other hand, involvement only varied with *Purpose* and not with the *Receiver*, which might suggest a lack of involvement in the relationship with the *Receiver* and a more task-oriented approach.

A further complication with the involvement and solidarity markers is their potential mutual exclusion. A sentence "We'll review this later" precludes the use of the involvement marker "I" as in "I'll review this later".

Finally the sociability marker did not vary with *Phase*, which may be attributable to its subcomponents. Different subcomponents of the social markers represent slightly different communication behaviours. Of particular note, is the difference between the trends of the general social building markers and politeness markers with time. Overall, general social

markers decrease with socialisation and politeness markers increase throughout the project. This highlights the different representations of component social markers, and the importance of understanding the situational representation of such markers. Politeness and courtesy may be erroneously interpreted as formality markers, but in this project their frequency increases with socialisation phase; i.e. with increasing familiarity individuals showed increasing courtesy.

Further research is needed to understand contextual definitions for coding, representations and interpretations of social markers in virtual writing teams, and I discuss this further in chapter 8.

5.4.7 Culture and language

The influence of culture and language is relevant in this study which includes two languages and an all Belgian team, except for one American. The first language of all the Belgian team members is Dutch, and the first language of the American team member is English. There were significant differences between these languages in formality measured by open greeting style and in sociability. One additional non-language related result, which may be attributable to culture was the high involvement shown by the American participant.

As one of the subcomponents of the social marker was politeness and courtesy, it occurred to me that these results may both represent higher formality in English emails, i.e. high open greeting formality and high politeness/courtesy formality. However, the main units counted for social markers were in the two categories general social building units (37) and politeness/courtesy (71). The ratios describing the relative numbers of these two types of units for both English and Dutch emails are 1.8 and 1.9 respectively (courtesy frequency: general social building units). Thus controlling for the difference in numbers of Dutch and English language emails, the relative numbers of the two types of markers in each language type were almost identical. The results here, therefore, show differences in two separate characteristics of language style in emails, formality and sociability. Thus, the English emails showed higher sociability, but the Dutch emails were less formal in open greetings.

The difference in formality measured by open greeting length reflects a difference in choice of communication behaviour, rather than differences in structure of the language for open greetings. This is clear from the first four most frequent open greeting behaviours, which were no greeting, name only and English and Dutch single word equivalents (see Figure 3-5). The influence of language and culture are confused in these emails, because some English language emails were written by Belgians. However, the fact that variations were apparent by *Language* for two of the communication behaviour markers shows that language and culture differences play a role in email style.

5.5 Chapter review

This study focused on the hypothesis:

H2 = Social dimensions of teams can be identified from email communications.

Indicators of influences on writing were derived from email records and shown to influence social behaviour in email communications in a networked team writing project. Writing influences analysed were *the Sender, Receiver, Audience* size, relational *Direction* of the email, *Purpose* and socialisation *Phase*. *Language* was also included as a control variable. The communication markers used to interpret social behaviour in team work were elaboration in body text and greetings of emails, style of greetings and use of first person pronouns and social building units. These markers were used to interpret effort and value attributed to communications, formality, involvement, solidarity and sociability. Participant feedback was used to help interpret the influences and identify task and social characteristics of the process.

At a macro-level, the results provide an overview of the task and socio-emotional components of email communications and thus a profile of the social behaviour which in itself contributes to team performance. This profile of social behaviour in terms of relational content and adaptations in relational metadiscourse, allows us to compare team writing projects and explore the influence of team culture on virtual team writing. All the writing influences showed significant variations in at least two of the communication behaviour markers. The most prevalent influence was *Purpose*, which affected all the communication behaviours, and represents the task dimension of the project.

Of surprising note in this study is that only three of the communication markers representing social behaviour, word count, formality and social markers, showed category differences for the writing variables by *Receiver*. Thus elaboration, formality and pro-social communication behaviours were used to adapt email style to help create a shared understanding between the writer and reader. Writers did not vary greeting style, involvement or solidarity to achieve a shared understanding with their *Receivers*. There are clear norms of behaviour in greeting behaviour (see Figure 3-5 and Figure 3-6), showing a social conformity which may have inhibited the team from adapting greetings extensively for an individual *Receiver*. There is also evidence to suggest that functional and organizational factors influence greeting style and involvement. The Kruskal-Wallis tests showed significant differences for both open and close greetings and involvement, for both *Direction* and *Purpose*. *Purpose* also affected adaptations in solidarity.

Involvement may reflect one or both of the task and social dimensions in the communication. Although writers adapted involvement with the *Audience* size, they also adapted involvement with *Purpose* and *Direction*, suggesting a tendency towards task involvement rather than socio-emotional involvement; involvement did not vary by *Receiver*, suggesting a lack of representation of the relational or socio-emotional element of the marker in this project.

Group cohesion increases to a peak at which performance is optimum, and then declines and a social dimension in team working contributes positively to the outcome (see section 2.3.2). The

balance between the social and task dimensions on a project therefore affects the team performance. The fact that there were only variations between *Receivers* for elaboration, formality and sociability, rather than for all of the communication markers suggests that the balance between social and task dimensions may lean towards the task dimension, with little socio-emotional communication.

The team members reported a general positive perception of group cohesiveness, although one team member had reservations. Greeting style, involvement and solidarity were not adapted for readers, and solidarity marker frequency was relatively low. While sociability was adapted for *Receivers*, one component of the social marker, which contributes to the building and maintenance of relationships, self-disclosure, was missing from the data. This profile of a low social to task balance in this project will be compared with an academic project in chapter 7.

At a micro-level, contextual information gathered from team members in pre- and post-analysis interviews, and combining evidence from multiple markers helps to add credibility to interpretations. Examples of interpretations which were reinforced by participant perceptions are those concerning the results for *Direction*. Profiles for elaboration by relational direction reflected the power hierarchy of the business relationship.

In terms of elaboration, the supplier made more effort and attributed more value for communications to the client, than for communications in house, and the client made less effort and attributed less value to the communications for the supplier than for communications in house. Profiles for elaboration by relational direction thus reflected the power hierarchy of the business relationship, and we see that the equalization effects predicted by Sproull and Kiesler's (1986) lack of social context cues hypothesis are not pervasive across email communication contexts, attributable in this case to the semi-virtual nature of the project.

There was also a relational difference in the tendency for higher formality in open and close greetings for the supplier to client emails than for the reverse emails. The results for *Direction* were also confirmed by the results by *Sender* for open greeting formality, with higher formality in supplier open greetings than client open greetings. Higher open greeting formality in the supplier context was explained by the interviewees as an organizational difference.

Involvement was higher on the supplier side than the client side, describing the supplier-client business relationship, which defines the supplier as the most active on the project. In house (as opposed to external) supplier interpersonal communications where relationships are likely to be more familiar also showed more involvement and less formal close greetings.

Thus the contextual relationship between writer and reader and organizational norms clearly influenced communication behaviour and email style in the team. This reflection of contextual relationships in communication behaviour demonstrates adaptation to the reader, albeit as defined by an organizational hierarchy, rather than the individual *Receiver*. Thus in this data, although writers appeared not to adapt greetings to *Receivers*, greetings did vary according to the writer's organizational context, and the organizational context of the *Receiver*, suggesting

the team demonstrated social interactive adaptations at an organizational level, rather than at an interpersonal level.

Although meaningful interpretations can be developed from situated knowledge and the email data, it is not possible to conclude from the study whether such interpretations at this micro level would be consistent for similar variations in different projects. To build a database of trends which can be interpreted consistently from empirical data extracted from emails would require similar detailed analyses on multiple projects.

Quantitative and qualitative data from this project suggest that the relationship between the social (or group maintenance) and task dimensions of the project is complex and cannot necessarily be analysed in a discrete way. For example, interviewees described reasons for expressing sociability and formality for particular *Purposes* in the project, thus using socio-emotional characteristics of communication for elements interpretively coded on the task dimension. Writers were more formal in communications on tasks completed early on e.g. *Accounts*, when they didn't know each other so well, and for purposes over which they felt more anxious, e.g. *Review.* Communications for *Courtesy*, which directly targeted group maintenance, were more sociable than other task oriented communications, but could also contribute towards functional goals. Task and social contributions at a team level are thus interdependent and cannot be analysed discretely, but rather need to be considered in a holistic way. The dual representations by markers, for example with involvement and elaboration representing both task and social components, and the potential masking of solidarity by use of involvement markers or vice versa, also necessitate a holistic interpretation. Additionally the fact that variations were apparent by *Language* for two of the communication behaviour markers shows that language and culture differences also play a role in email style.

Showing team members (albeit preliminary) analyses of the data in the post-analysis interviews encouraged them to search their own situated knowledge of the project, often resulting in logical explanations for the trends they observed, such as changes in solidarity levels when the project leader changed. This member-checking technique for analyses of email data may be valuable in management practice as a corrective intervention in team working, to help team members understand and improve the social dynamics of the project.

In conclusion, I have shown in this study that social dimensions of teams can be identified from email communications, and that this information helps to describe the balance of task and social dimensions in networked team writing projects. The results at a macro level, considering the behaviours adapted by writers to achieve a shared understanding in both task and social components of communications, contribute to provide a holistic overview of team behaviour. The results provide an overview of the task and socio-emotional components of email communications and thus a profile of the social behaviour which in itself contributes to team performance. This profile of social behaviour in terms of relational content and adaptations in relational metadiscourse, allows us to compare team writing projects and explore team behaviours which may influence virtual team writing. I explore this further in chapter 7.

6. H2 Academic case study: team social dimensions

6.1 Research focus

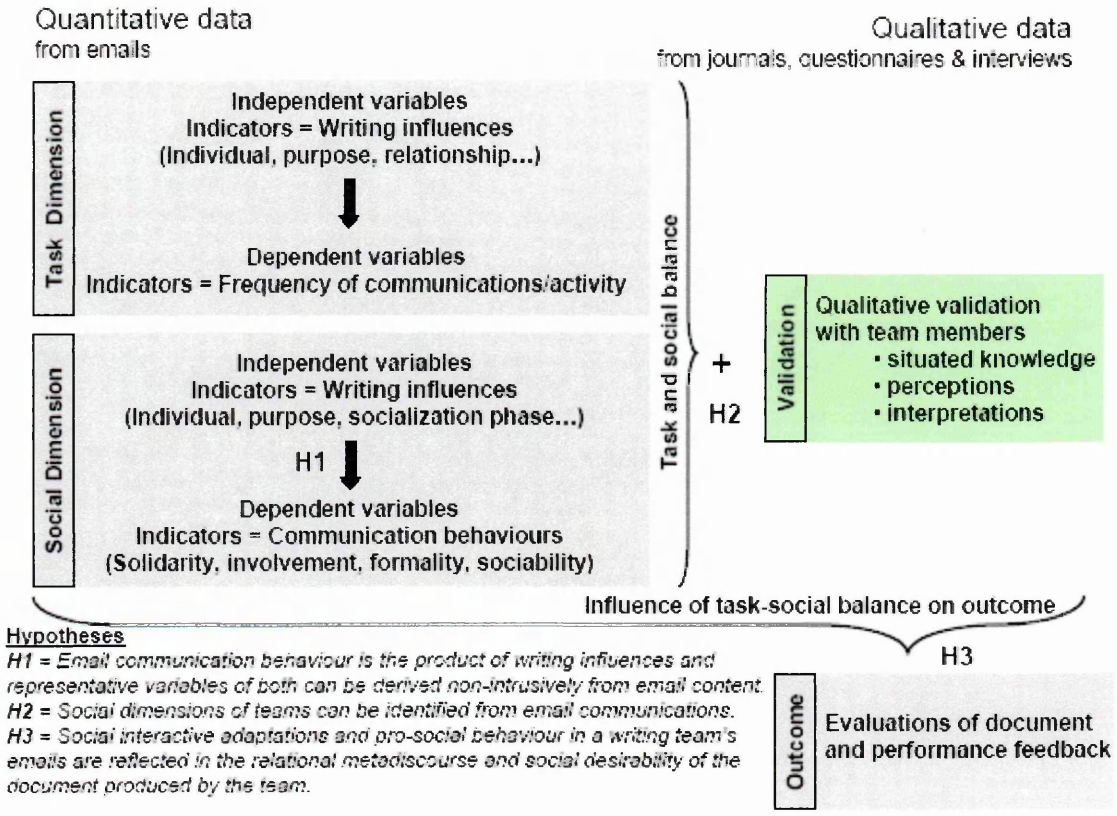

Figure 6-1: Research Framework highlighting H2 phase

This study explores interpersonal email communications during a professional networked team writing project to develop a post-graduate training course. It is the third study in a series of analyses to support the development of an email analysis tool, which allows non-intrusive research into writing practice and comparison of different writing projects in a consistent way.

The previous studies reported in chapters 4 and 5 have showed that indicators of influences on writing could be derived from email records and shown to influence social behaviour in email communications in networked team writing projects. This study applies the same process to test whether dependencies occur in an academic context, and to interpret these in the light of the main coordinator's feedback. This part of the research is shaded green in Figure 6-1, which shows how this study fits within the overall research framework and addresses the following hypothesis:

H2 = Social dimensions of teams can be identified from email communications.

866 emails exchanged between the course coordinator and main team members between February 2000 and May 2005 provide a corpus of communications representing the Society

discourse during the course development. Email contents were coded to extract the following writing influences, the independent variables, *Sender, Receiver, Audience size, Direction, Purpose, Phase* and *Language*. The following communication behaviour markers, the dependent variables, were also derived from the email content and coding: word count for body text, open and close greetings, % first person singular and plural pronouns, formality score and frequency of social building units. Empirical data derived from the coding were analysed using SAS statistical analysis package. Pre and post analysis interviews were completed with the main coordinator of the postgraduate course to collect information on the project and the team members, to help develop interpretations from the email data. For a full description of the methods used in this analysis and the background to the Society project, please see chapter 3.

In section 6.2, I present the results of the data analysis. Significant results for each communication marker are then analysed in depth in section 6.3 and triangulated with interviewee ratings for social behaviour of team members, to develop interpretations.

6.2 Results: significant writing influences

The Kruskal-Wallis computation tests the null hypothesis that categorical variables (such as *Sender* in this study) have identical distribution functions against the alternative hypothesis that at least two categories differ with respect to location. If the p value is small, we can reject the idea that the difference is a coincidence, and conclude instead that the populations vary. The appropriateness of this test for the data to be analysed is discussed in section 3.7.3. An extract from the SAS output for Kruskal-Wallis test for the independent variable, *Sender*, and dependent variable social markers is presented in appendix JJ.

Table 6-1 shows the p values for the Chi square value resulting from the Kruskal-Wallis tests for each paired combination of independent and dependent variable, i.e. for each paired combination of a writing influence and email communication behaviour.

Table 6-1: p values for Kruskal-Wallis Chi square values

(Shaded cells indicate p is significant at the 0.05 level)

Pr > Chi square		Dependent variables: communication markers						
		Word count	Open greeting	Close greeting	% first person singular pronouns	% first person plural pronouns	Formality	Social building units
Independent Variables: writing influences	Sender	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
	Receiver	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
	Audience	.1585	<.0001	.0841	<.0001	.0009	<.0001	<.0001
	Direction	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
	Purpose	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
	Phase	<.0001	<.0001	<.0001	<.0001	<.0001	.0076	.0001
	Language	<.0001	.8453	.2331	.6495	.0078	.0429	.0400

All the writing influences studied caused significant variations in the communication behaviours with the following exceptions: email body text length (word count) and close greeting length did not vary with *Audience* size, and open and close greetings and involvement markers (% first person singular pronouns) did not vary with *Language*. The Kruskal-Wallis statistic only tests whether there are overall differences between the categories of a variable, but not the nature of the differences (Greene and D'Oliveira 2006 p79). To interpret the nature of the differences between the categories requires interpretation of the trends shown by descriptive statistics of the communication behaviour variables and these are discussed further in the following sections.

6.3 Interpretations of social behaviour

6.3.1 Value attributed to communications: elaboration

Elaboration in email body texts

Elaboration represents effort and value attributed to a communication, either over task or socio-emotional content. Elaboration showed significant differences between the distributions of at least two categories of all writing influences (p= < 0.0001) except for *Audience size*.

High elaborators by *Sender* are AF1 and AF10 (see Figure 6-2). These team members were responsible for 4% and 3% of the total email communications respectively (email frequencies by *Sender* are shown in appendix K line 220). Low elaborators are the President, the Administrator, AF7 and AF13, who were responsible for 8%, 20%, 4% and 1% of the total email communications respectively. Elaboration and email frequency data thus show that effort and value assigned to a communication is partly individualistic and independent of the overall activity contributed by an individual to the project.

Figure 6-3 shows how writers adapt the effort and value in a communication for the *Receiver*. For example, emails addressed to the President, to AF9, to Academic Faculty members (AF Dist. list) and to mailing lists (also subject matter experts) had higher elaboration. Emails circulated between members of the conference organizers (Organizer distribution list) had lower elaboration. Thus writers adapt how much effort is attributed to the communication according to the *Receiver(s)*.

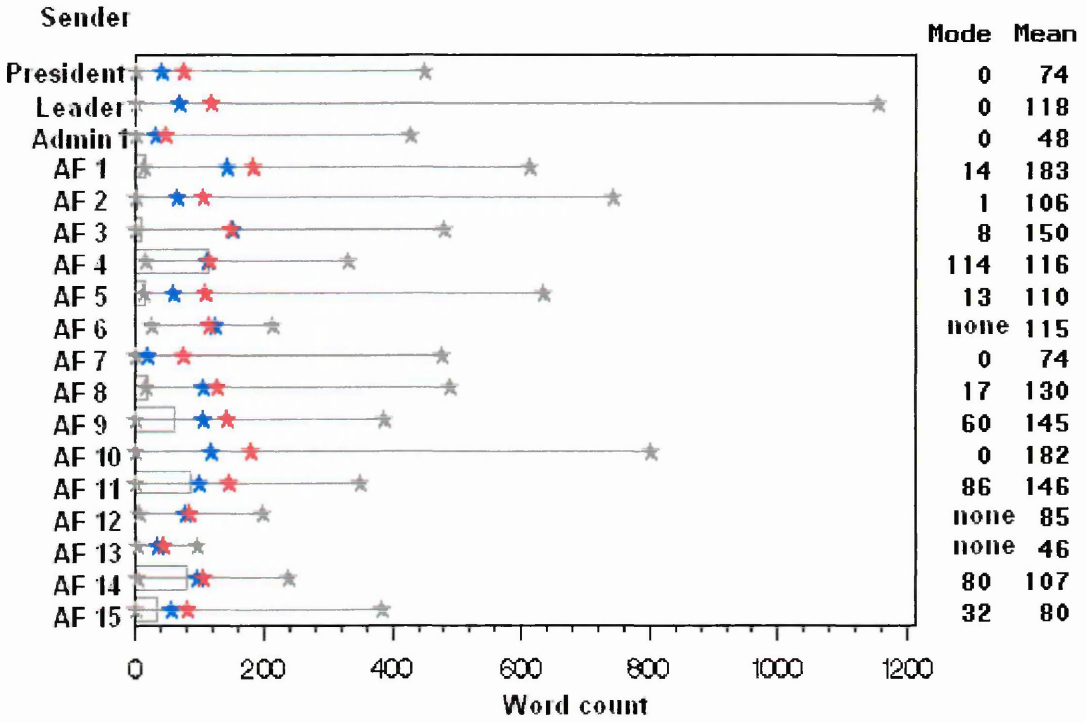

Figure 6-2: Differences by sender for body text word count

Legend: min max median mean mode

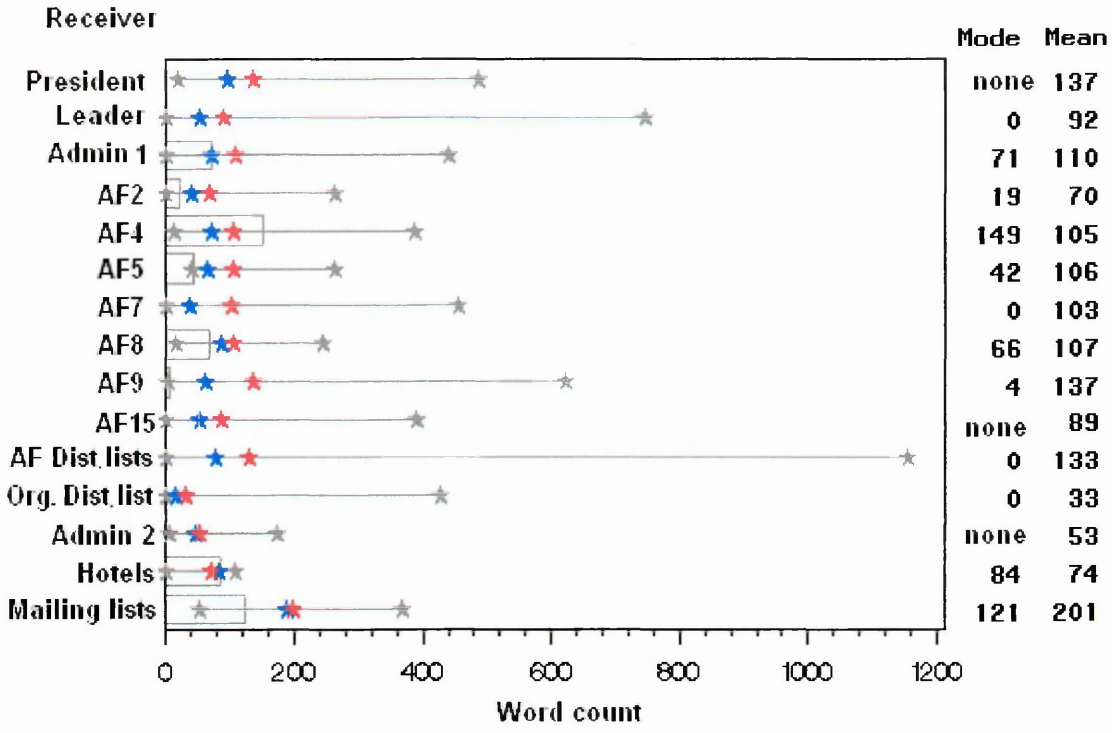

Figure 6-3: Differences by receiver for body text word count

Legend: min ★ max ★ median ★ mean ★ mode
Admin = Conference administrator AF = Academic Faculty members
Leader = Course Leader; P = President; Org. Dist. List = Organizer's Distribution List

Highest elaborations by direction (see Figure 6-4) were between the Course Leader and President (CL→P) and between members of the Academic Faculty (AF→AF). Lowest elaborations were between the Course Leader and Administrator and vice versa. This reflects the high need for elaboration over complex subject matter content, and lower need for elaboration over less complex administrative and organizational issues. This interpretation is also substantiated by the higher elaborations by *Purpose* for *Review* and *Product* discussions compared to *Management* and scheduling as shown in Figure 6-5. The *Product* in this study is the actual course content. *Purposes* with low elaboration were *Accounts*, *Management*, *Courtesy*, *Document* transfer, *Circulation* of information and *Technology* issues. The *Technology* category covered application and mediation difficulties, such as when individuals could not open or send attachments. Elaboration is therefore higher for *Purposes* involving discussion over the Society's subject matter expertise, and lower when emails are used as transfer agents or for group maintenance issues such as *Management* and *Courtesy* emails.

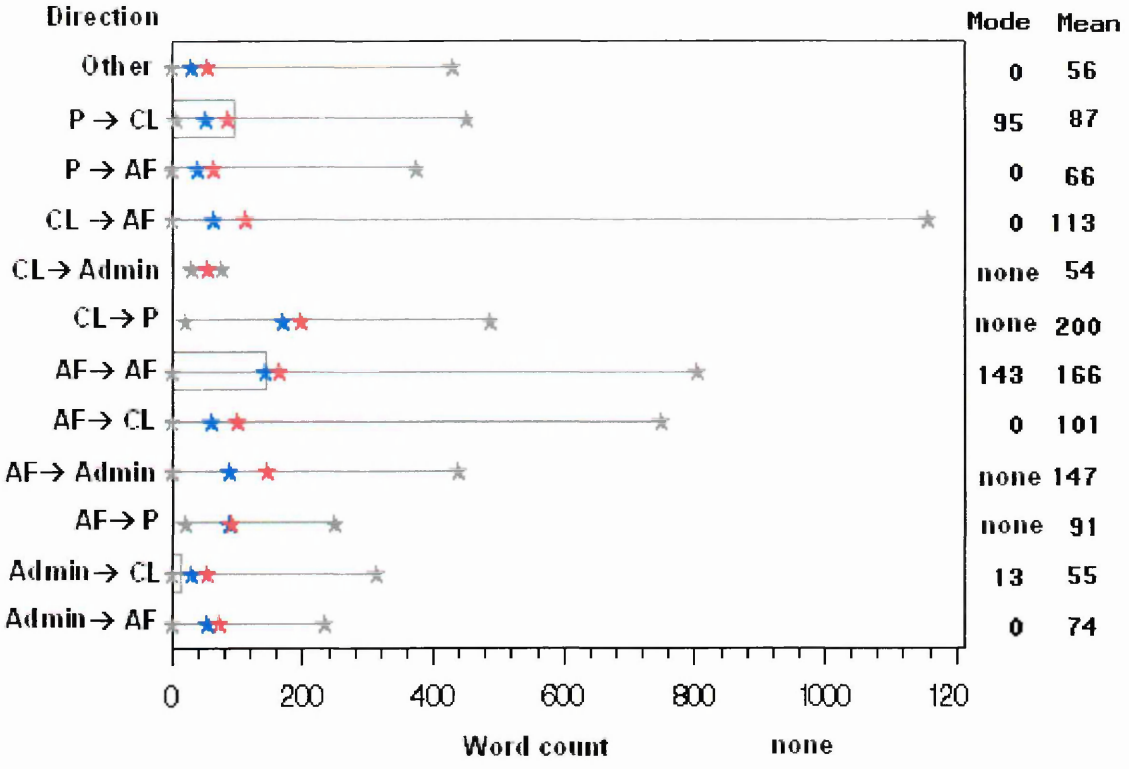

Figure 6-4: Differences by direction for body text word count

Legend: min ★— max —★ median ★ mean ★ mode ▭

Admin = Conference administrator AF = Academic Faculty members

CL = Course leader P = President

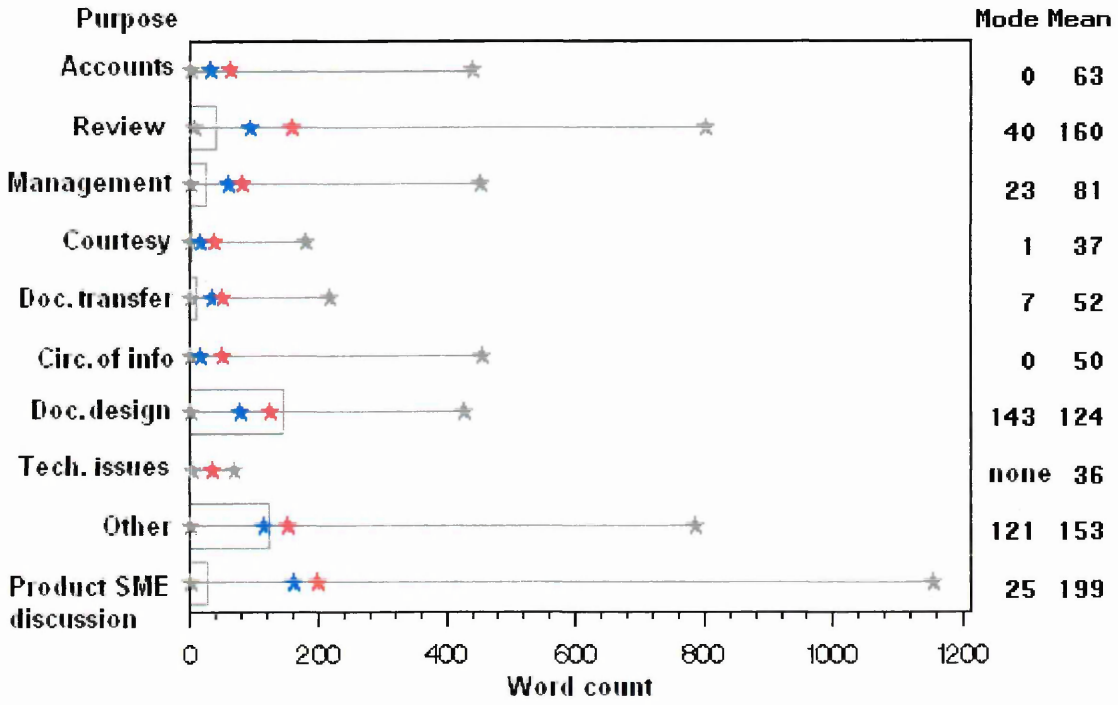

Figure 6-5: Differences by purpose for body text word count

Legend: min ★— max —★ median ★ mean ★ mode ▭

The differences in elaboration by *Phase* shown in Figure 6-6 show higher initial elaboration which then falls during phases two and three. There is then a change between phase three and four; phase three ended September 2002. Interestingly, in the interview with the Course Leader following a preliminary analysis of the data, the interviewee twice referred to changes after the 2002 course and conference, and one of her comments pertains to the task dimension:

> *...there may have been a slight change after the 2002 seminars took place – evaluation and reassessment of aims and goals, around October 2002, but the group hasn't changed (Society interviewee: appendix K line 72).*

A plot of *Purpose* frequencies by *Phase* (see appendix KK) shows an increase in phase four for both *Product* and *Review* discussions substantiating the interviewee's comments. Thus the increase in elaboration after the third socialisation phase of the project reflects the renewed effort addressed towards the task, expended in terms of review of the course performance.

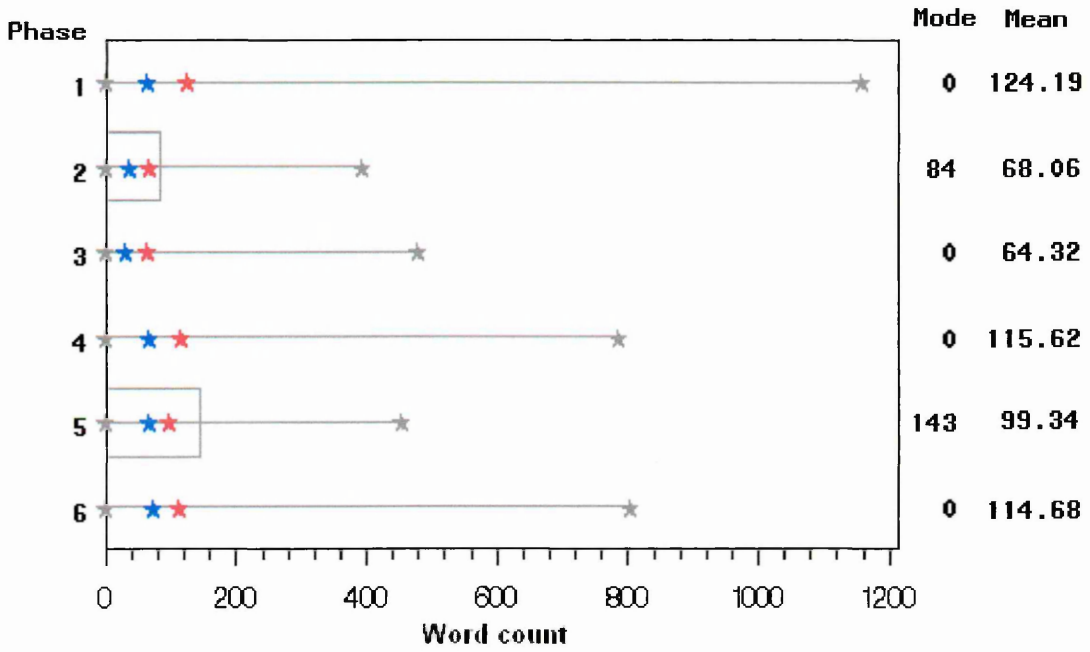

Figure 6-6: Differences by phase for body text word count

Legend: min ★— max —★ median ★ mean ★ mode ▭

There was also a significant difference between elaboration in English and Dutch language emails, with Dutch emails having higher average word counts (see Figure 6-7). This suggests that emails sent by Dutch speaking *Senders* to colleagues in Belgium or the Netherlands were longer than those sent to colleagues elsewhere. There were four Dutch speaking Senders (see appendix K line 45-46). Two of the Dutch speaking team members were low elaborators (Admin1 and AF7) and two relatively high elaborators (Course leader and AF3 in Figure 6-2).

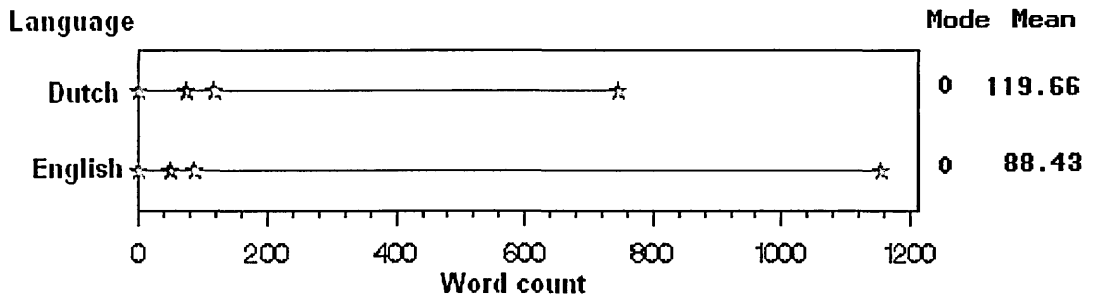

Figure 6-7: Differences by language for body text word count

Legend: min ☆— max—☆ median ★ mean ☆ mode ▭

Elaborations in greetings

Greeting lengths were used in an earlier study to represent formality (see section 3.7.2, Figure 3-5 and Figure 3-6). The frequencies of greetings types in this study have been presented in chapter 3 (Figure 3-9 and Figure 3-10), and again in this data suggest a trend of increasing formality with increasing word count for the most frequently used greetings. There is a dual representation in greetings, therefore, of both formality and effort and value attributed to the *Sender-Receiver* relationship.

The Kruskal-Wallis test showed that open greeting lengths varied between at least two categories of all the writing influence variables (p= <0.0001) except for *Language*. Close greeting length varied significantly between at least two categories of all the writing influences (p=<0.0001) except for *Language* and *Audience* size.

Variations by *Sender* for open greeting can be seen in Figure 6-8 and by *Sender* for close greeting in Figure 6-9. We can see from the values for modes that there are three most frequent open greeting styles, zero, one, and two words; individuals vary in the style which they most frequently use for both open and close greetings. However, *Senders* who use a shorter (less formal) open greeting do not necessarily also use a shorter close greeting. For example, the Course Leader and AF2 both use two words in open greetings, but include no close greeting. Thus it is clear that style of open and close greeting varies by individual and that individuals may have different open and close greeting styles.

Additionally, out of the three Senders with the shortest open greetings, AF4, AF7 and AF13, two (AF4 and AF13) were rated by the interviewee as relatively formal, and only AF7 was rated as relatively informal. Examples of *Senders* with long close greetings are the President, AF8 and AF15; the President and AF8 were rated as relatively formal and AF15 as middle of the range by the Society interviewee (see Table 6-3 or appendix K, line 168). Inconsistencies between interviewee ratings and formality interpreted from greeting lengths highlight the dual representation of greeting lengths as both markers of formality and markers of effort and value attributed to the *Sender-Receiver* relationship. Short open greeting styles may represent relatively low formality or relatively low effort and value attributed to the communication or to the relationship with the *Receiver*.

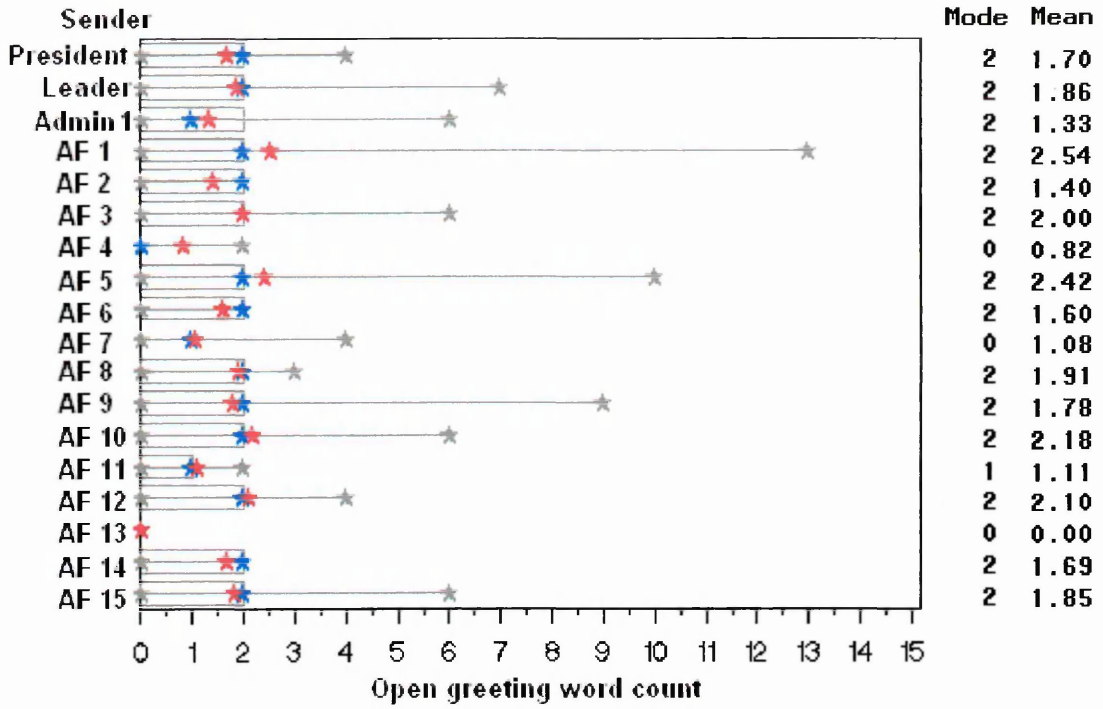

Figure 6-8: Differences by sender for open greeting word count

Legend: min ★— max —★ median ★ mean ★ mode □

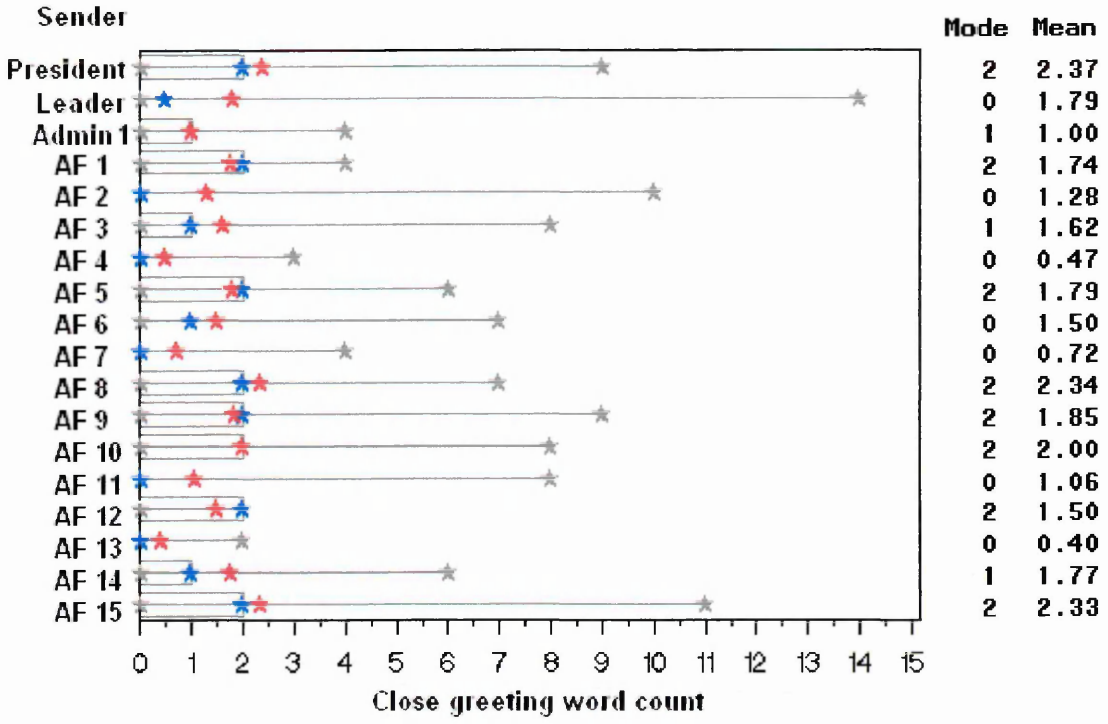

Figure 6-9: Differences by sender for close greeting word count

Legend: min ★— max —★ median ★ mean ★ mode □

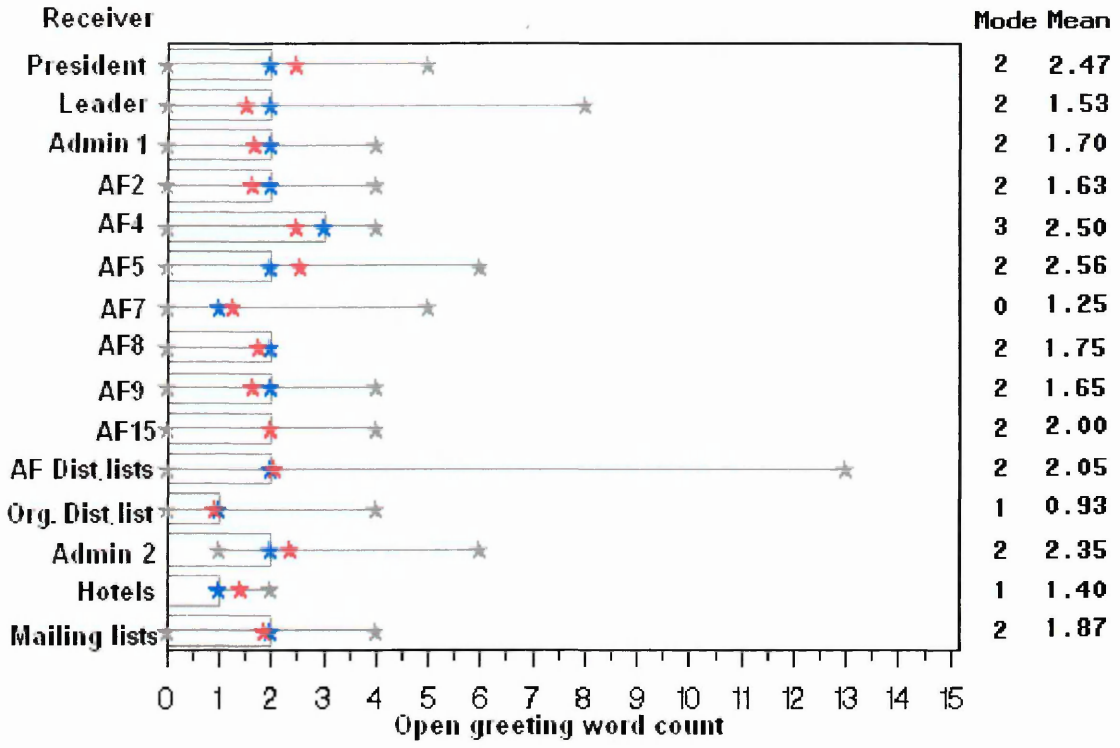

Figure 6-10: Differences by *Receiver* for open greeting word count

Legend: min ★ max ★ median ★ mean ★ mode
Admin = Conference administrator AF = Academic Faculty members
Leader = Course Leader; P = President; Org. Dist. List = Organizer's Distribution List

Longest open greetings by *Receiver* (Figure 6-10) were written for the President, AF4 and AF5, who were all rated relatively formal by the Society interviewee (See Table 6-3 or appendix K, line 168). Shortest greetings were written for AF7, who was also rated as least formal and who had relatively high FtF contact, and for members of the organizer's distribution list, a small team of three who organized the 2002 conference, and who had the most FtF contact (appendix K lines 62 and 175).

Relatively shorter open greetings were also used in the emails to hotels, which also had the longest close greetings (Figure 6-11). These emails were enquiring about accommodation, i.e. business communications between a potential client and supplier. This example illustrates that adaptation of open greeting length for *Receivers* is not necessarily consistent with adaptation of close greeting length.

Receivers who most frequently received no close greetings were AF2, AF7, AF8 and AF9, all of whom were rated as relatively informal except for AF8. Interestingly, *Receiver* AF8 actually shows a relatively high mean for close greeting length, although the mode was 0.

Thus there are examples of adaptations in open and close greetings for *Receivers* which matched the interviewee's formality ratings, demonstrating writers' adaptation of these markers to the formality of the reader. However, variations in style of open and close greeting suggest that the markers represent different socio-emotional components of email communication.

Additionally, the low greeting lengths sent to AF7 and the Organizer's distribution list, include the influence of FtF contact on email style.

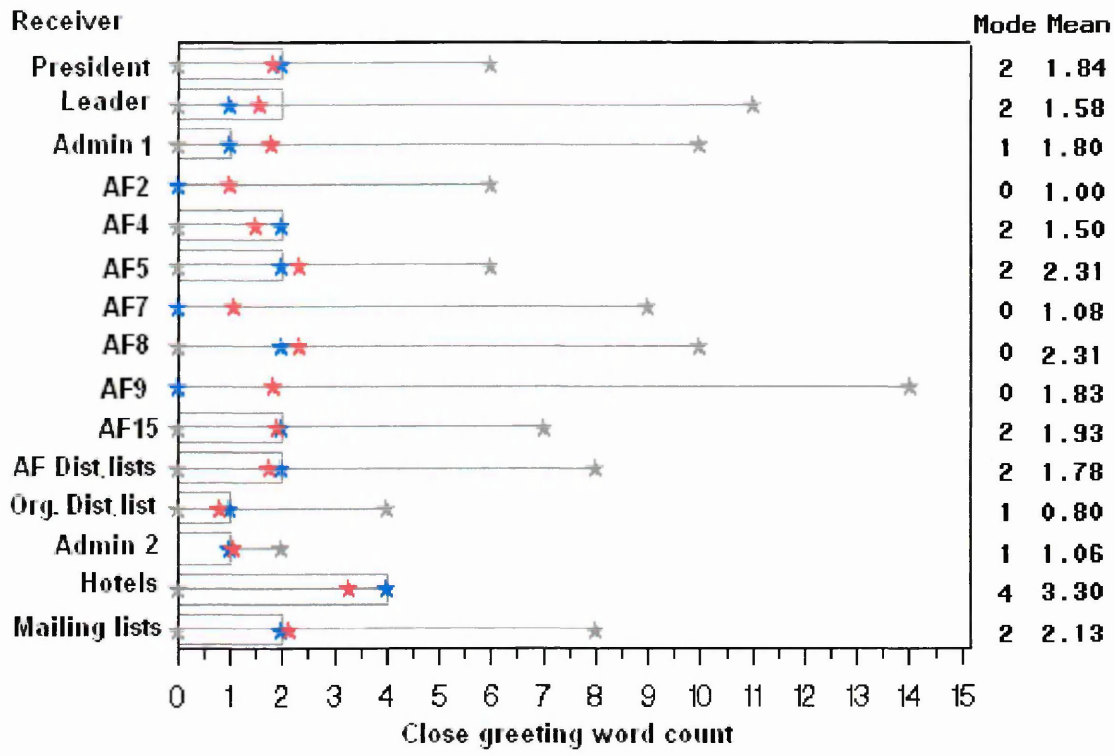

Figure 6-11: Differences by *Receiver* for close greeting word count

Legend: min ★— max —★ median ★ mean ★ mode ▭

Admin = Conference administrator AF = Academic Faculty members

Leader = Course Leader; P = President; Org. Dist. List = Organizer's Distribution List

Mean open greeting length increased, becoming more formal, with increasing audience size (see Figure 6-12).

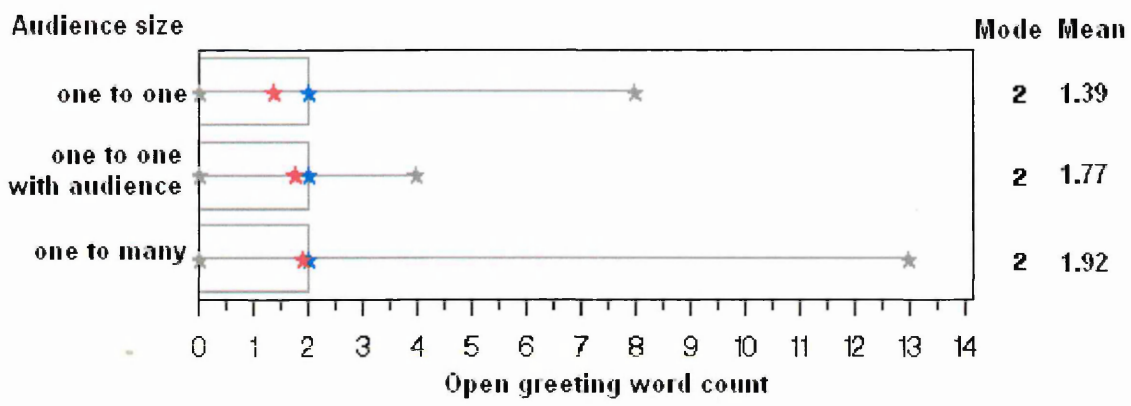

Figure 6-12: Differences by audience size for open greeting word count

Legend: min ★— max —★ median ★ mean ★ mode ▭

Similar to elaboration measured from body word count, mean elaboration in open greetings by direction (Figure 6-13) was highest between the Course Leader and President.

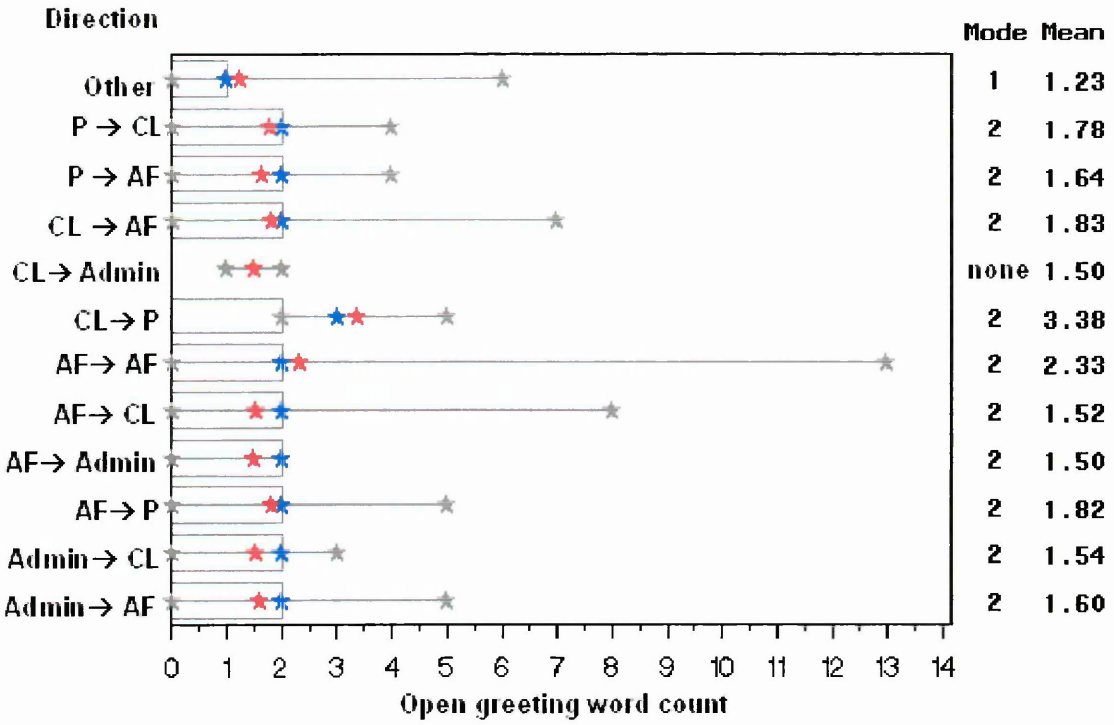

Figure 6-13: Differences by direction for open greeting word count

Legend: min ★ max ★ median ★ mean ★ mode
Admin = Conference administrator AF = Academic Faculty members
CL = Course leader P = President

There was a tendency for close greetings to or from the Administrator to be shorter, and therefore less formal than others, (see Figure 6-14); the most frequent behaviour by the Course Leader was omission of close greetings to all addressees, although mean close greeting lengths were higher in emails to the Academic Faculty and President than in those to the Administrator.

Mean open greetings were longest for *Review* and *Product* discussions, and lowest for *Courtesy* messages and *Circulation* of information (see Figure 6-15). Close greetings showed a similar trend with highest means for *Review* and *Product* discussions (see Figure 6-16) and short close greetings for *Circulation* of information.

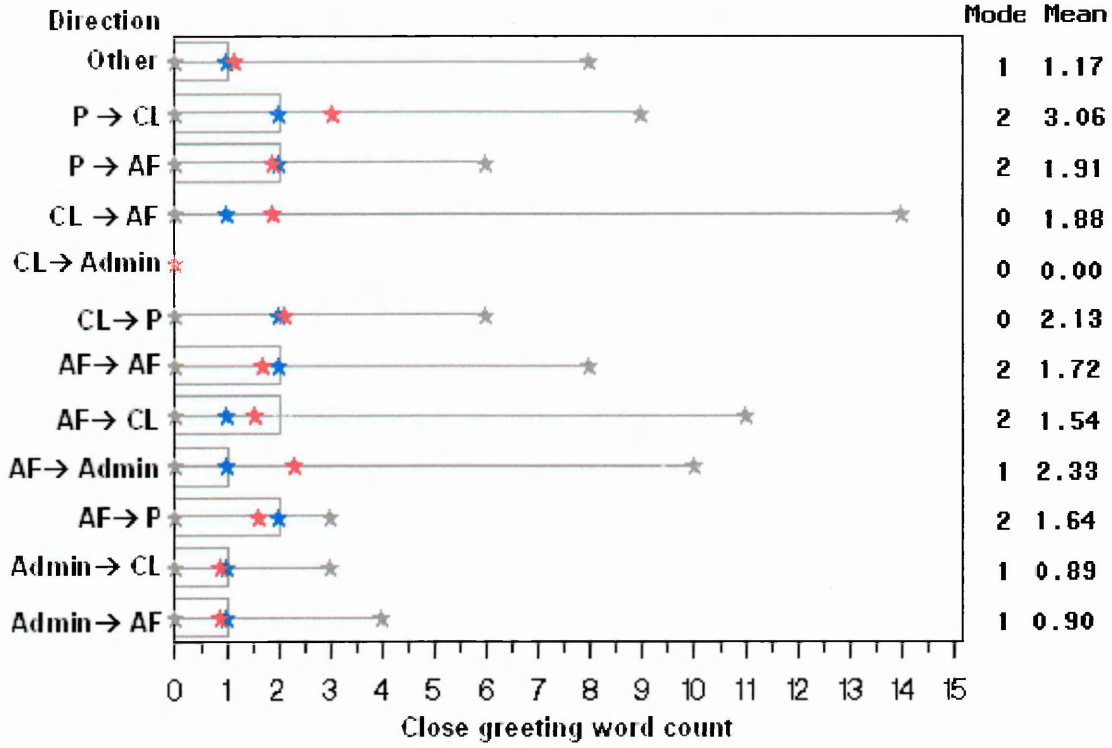

Figure 6-14: Differences by direction for close greeting word count

Legend: min max median mean mode

Admin = Conference administrator AF = Academic Faculty members
CL = Course leader P = President

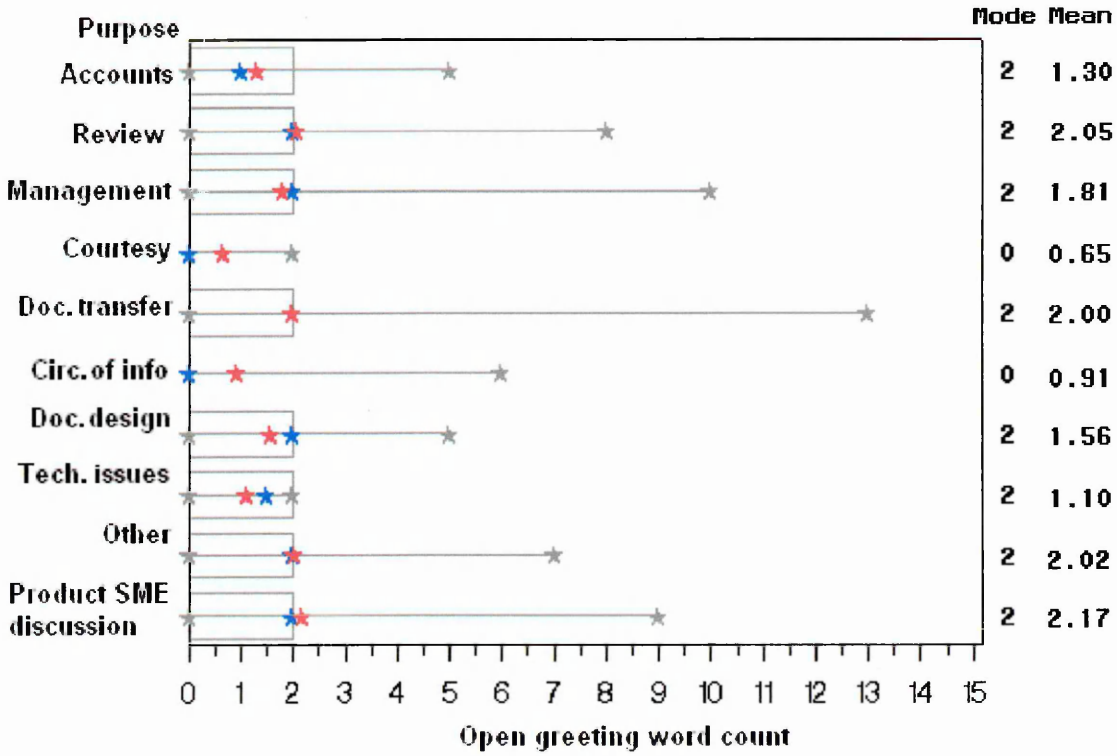

Figure 6-15: Differences by purpose for open greeting word count

Legend: min max median mean mode

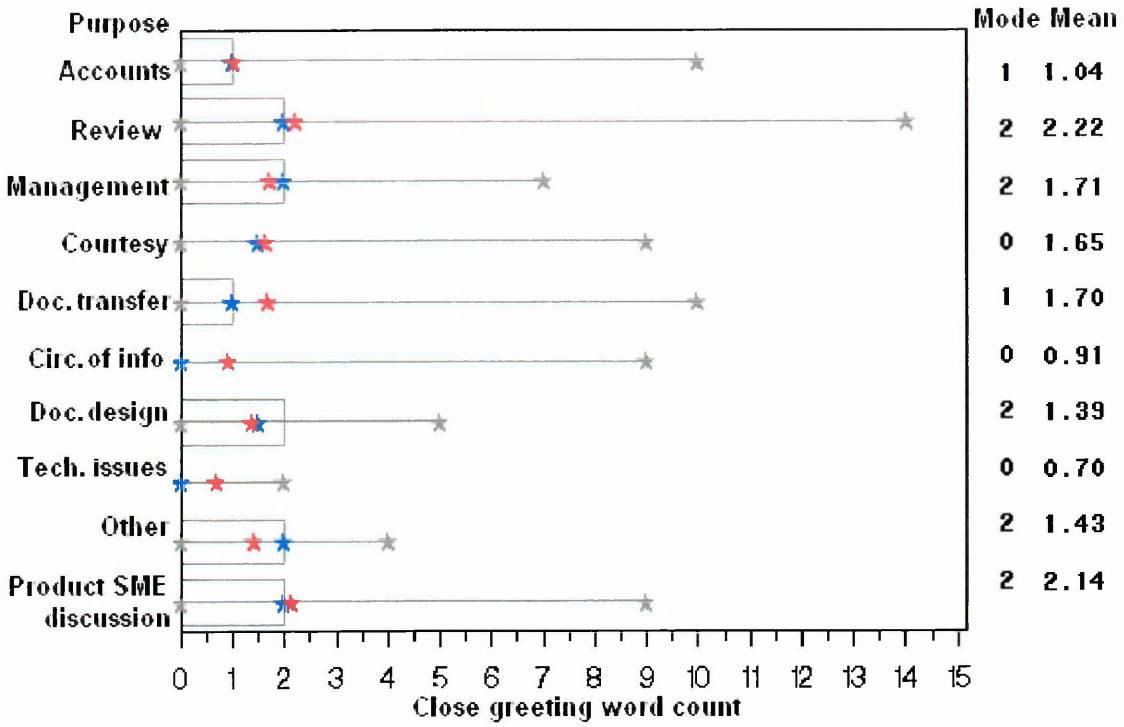

Figure 6-16: Differences by purpose for close greeting word count

Legend: min ★ max ★ median ★ mean ★ mode

Finally elaboration in both open and close greeting showed a similar trend to elaboration in body text, with a fall during the first three phases, and then an increase in effort and value attributed to messages in phase four, following the 2002 conference.

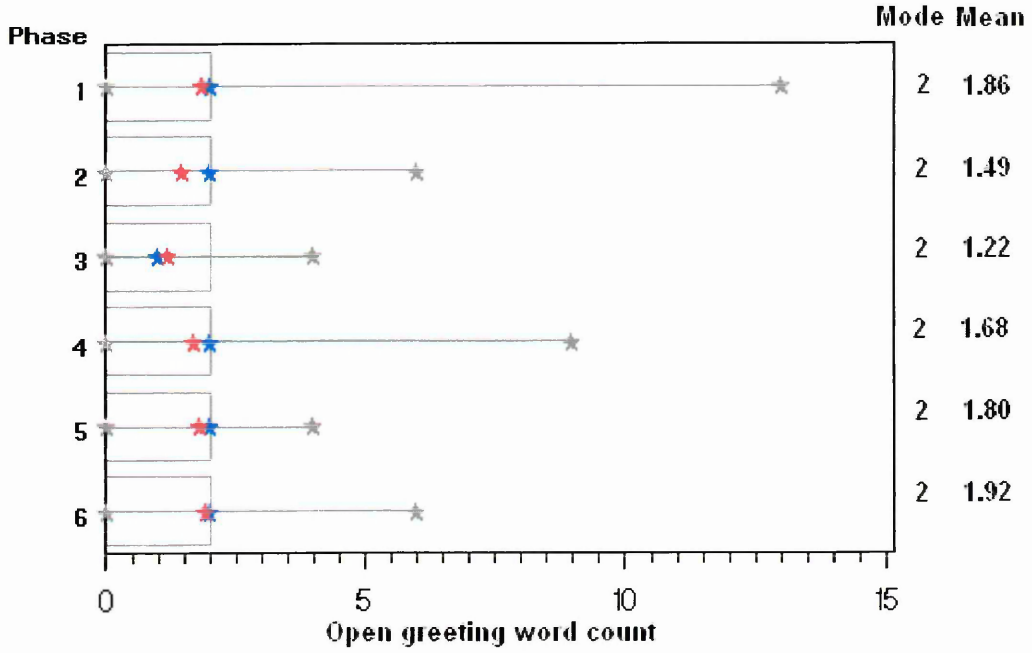

Figure 6-17: Differences by phase for open greeting word count

Legend: min ★ max ★ median ★ mean ★ mode

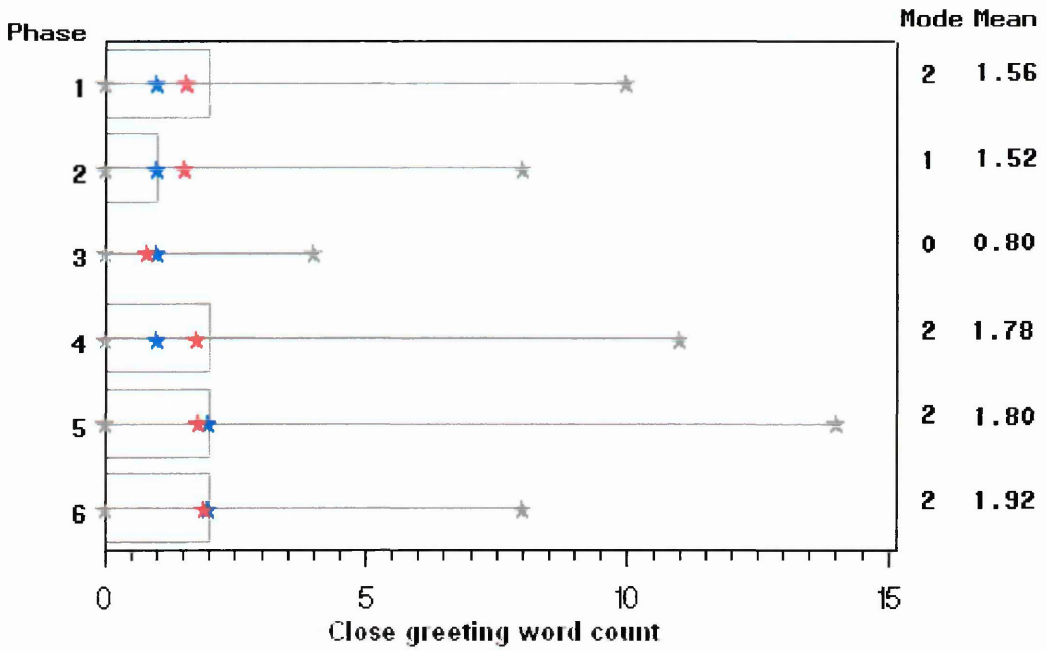

Figure 6-18: Differences by phase for close greeting word count

Legend: min ★— max —★ median ★ mean ★ mode □

6.3.2 Involvement interpreted from first person singular pronouns

First person pronouns are used in this study as markers of involvement in either the socio-emotional or task dimensions. Frequency of the involvement marker was higher than the solidarity marker (involvement total count 3121, solidarity total count 1072 in 866 emails). Involvement varied between at least two categories of all the writing influences ($p = < 0.0001$) except for *Language*.

The interviewee's ratings for formality by team member are presented in Table 6-2. Analysing the interviewee's ratings showed a high correlation between her ratings for activity on the project and ratings for involvement ($r = +0.8$ $p = 0.0004$), sociability ($r = +0.5$ $p = 0.0327$) and formality ($r = -0.8$ $p = 0.0005$) (see appendix LL). Thus the interviewee tended to rate team members who she perceived as more active on the project, as also more involved, more sociable and less formal.

Table 6-2: Interviewee involvement ratings, sorted by rating

Involvement (1=most involved)	Team member
1	Course Leader
2	AF2
3	AF9
4	AF14
5	Administrator
6	AF7
6	AF12
7	AF10
8	AF1
8	AF11
9	AF5
9	AF6
10	President
10	AF3
10	AF4
10	AF8
10	AF13
10	AF15

Figure 6-19 shows the involvement represented in emails by first person markers. Two individuals showing relatively high involvement are AF13 and AF14. Interestingly, the interviewee rated AF14 as the fourth most involved, but only seventh most active in the project (appendix K, line 45-46). The interviewee's judgement on involvement for AF14 is therefore reflected in the involvement markers from the emails. Also, activity represented by email frequency for AF14 was only 1% (appendix K, line 220), which conforms to the interviewee's perception of this individual's activity on the project. This example illustrates the qualitative nature of the data extraction, in that email frequency shows relatively low activity for this individual, but the email content shows high involvement in the activities contributed, also validated by the interviewee. Involvement and activity markers thus represent different aspects of an individual's contribution towards the project.

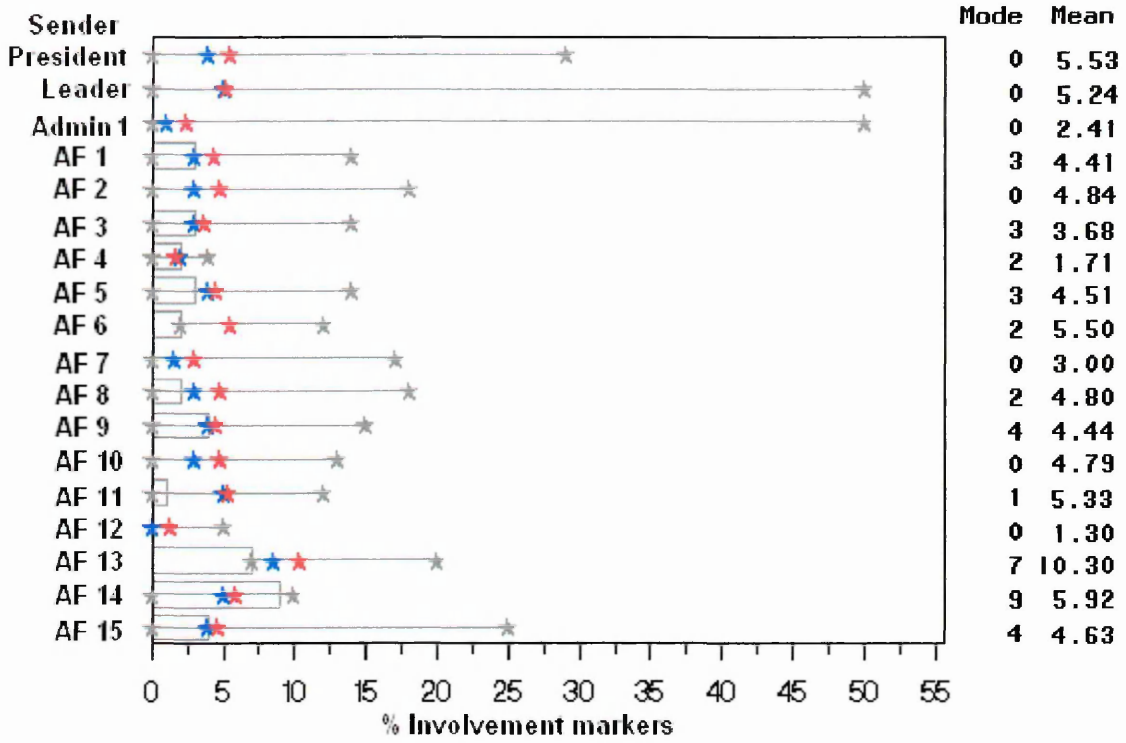

Figure 6-19: Differences by sender for involvement

Legend: min max median mean mode

However, the interviewee's feedback for AF13 does not conform to the email data. This individual was rated by the interviewee with relatively low involvement and activity (see Table 6-2 and appendix K, line 45-46, 167). The interviewee feedback, therefore, does not explain the high frequency of involvement markers for this individual. Example texts from the emails written by AF13, however, justify the interpretation of high task involvement:

> *Thank you for your invitation. I am pleased to accept. I will send the CV and photo later- remind me if I forget (Text 3188).*
>
> *I am at a conference next week but I will send what you need when I get back (Text 6766).*
>
> *I hope these are adequate (Text 3036).*
>
> *I do not intend to bring my laptop so I can either bring slides or a CD. Will it be OK if I email the handout sheets next week (black & white) or shall I send a disc? (Text 3001).*
>
> *I have been trying all day to email the PowerPoint file for the handouts and I keep reducing it to make it smaller but still it won't download. I will try again now (Text 2985).*

AF13 appears to be involved in the task in the excerpts above, but only contributed 1% in terms of email communications (appendix K, line 220), and was rated as one of the less active on the project (appendix K, line 45-46). It is possible, therefore, that whereas AF13 did not contribute largely to the project, her task involvement was high in what she did contribute, and this is reflected in the involvement markers. As the interviewee's perceptions of involvement correlated

highly with her perception of relative activity, low relative activity may have biased her perception of this team member's involvement.

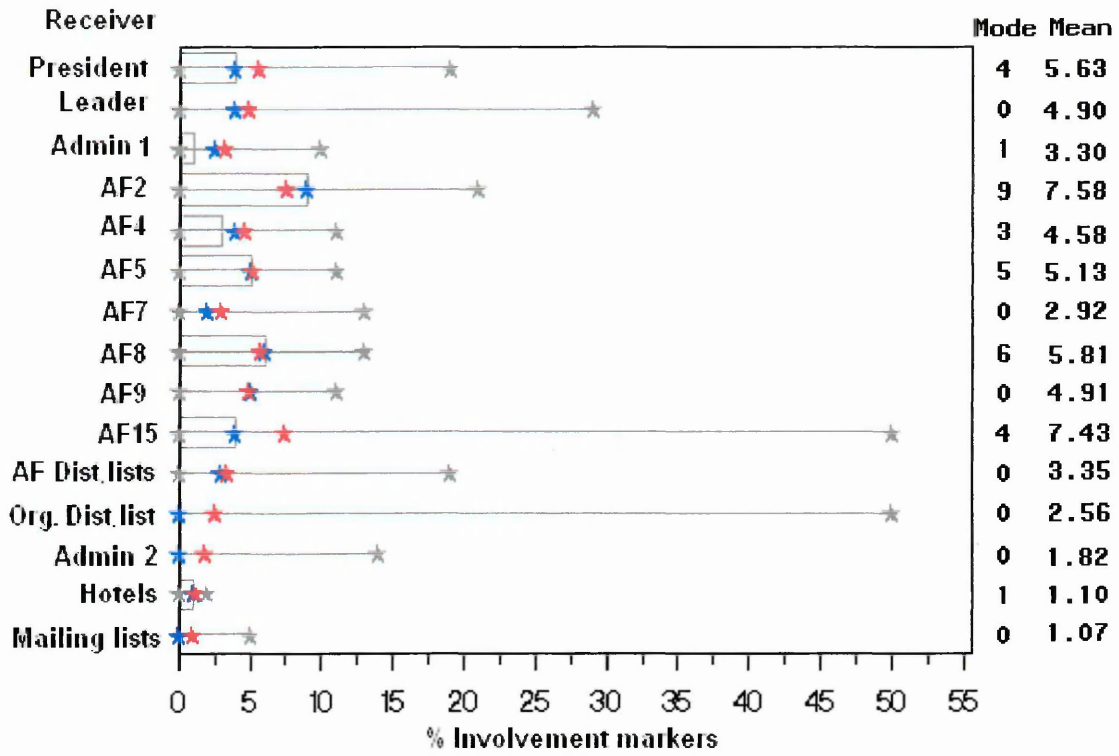

Figure 6-20: Differences by receiver involvement

Legend: min max median mean mode
Admin = Conference administrator AF = Academic Faculty members
Leader = Course Leader; P = President; Org. Dist. List = Organizer's Distribution List

Writers adapted their email style in terms of representation of involvement according to who they were writing to. Figure 6-20 shows that in particular emails to AF2 had more involvement markers, and this individual was rated by the interviewee with relatively high involvement. Thus in this example, writers adjusted their own representation of involvement to match involvement of the *Receiver*. Involvement markers are lowest for the two categories of *Receiver* which are most likely to be unknown to the writer, mailing lists and hotel personnel.

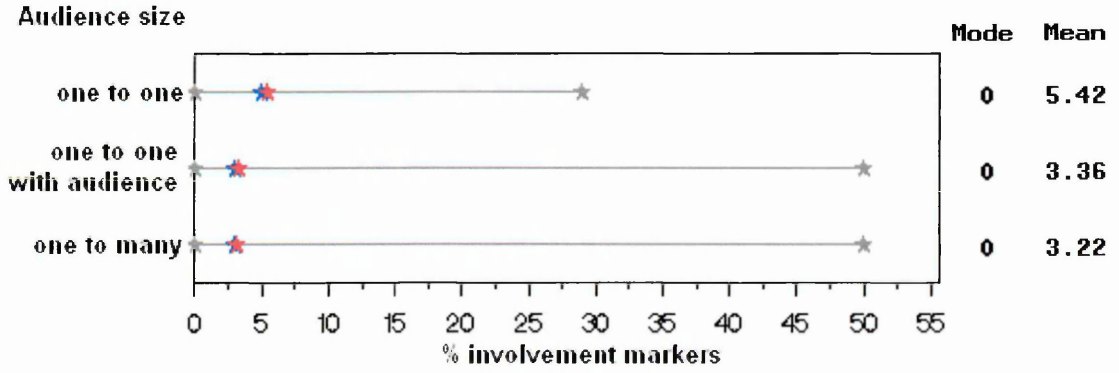

Figure 6-21: Differences by audience size for involvement

Legend: min max median mean mode

Figure 6-21 shows decreasing involvement with increasing audience size. This reflects the socio-emotional dimension of involvement, with higher involvement for private interpersonal communication.

Highest involvement by direction is represented in emails from the Course Leader to the President and vice versa, and from the Course Leader to the Administrator. The two most active members of the team in terms of email frequency were the administrator (20%) and the Course Leader (18%). The interviewee also rated herself and the Administrator as the two most active on the project (appendix K, line 45-46), and the Course Leader by functional role should be the most involved in the task, and indeed, relatively high task involvement is reflected in all the communications from the Course Leader (see Figure 6-22). The high involvement in communications from the President to the Course Leader also reflects the President's involvement in the project (8% and third most active in terms of email communications).

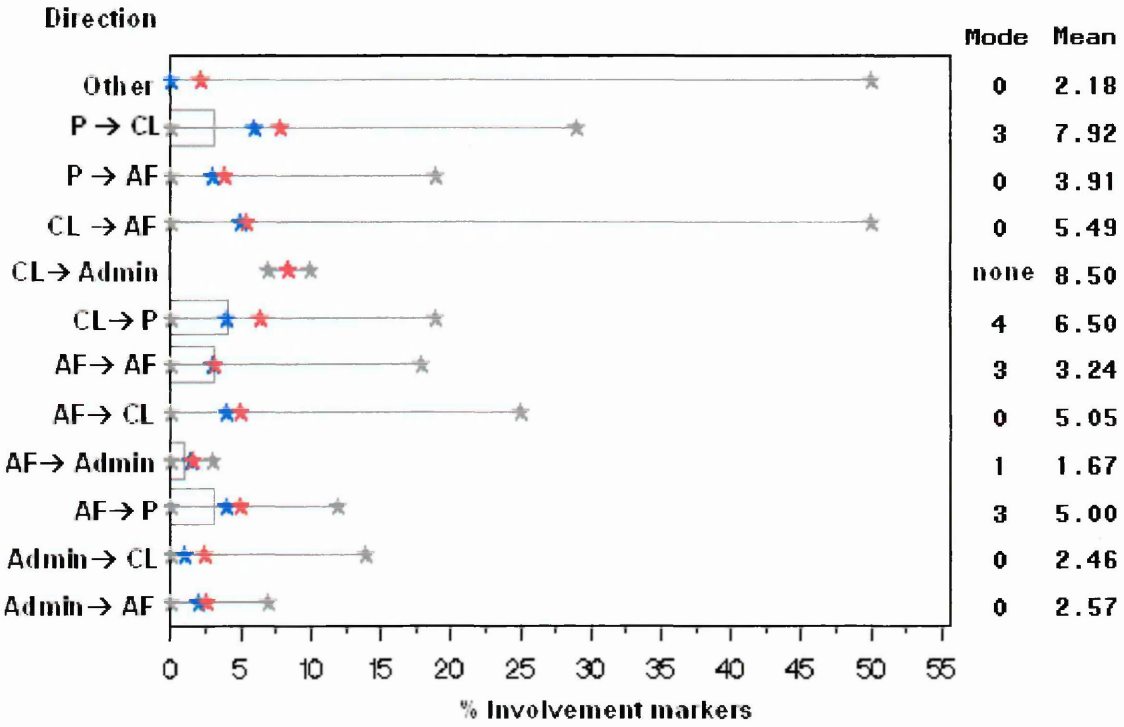

Figure 6-22: Differences by direction for involvement

Legend: min ★ max ★ median ★ mean ★ mode
Admin = Conference administrator AF = Academic Faculty members
CL = Course leader P = President

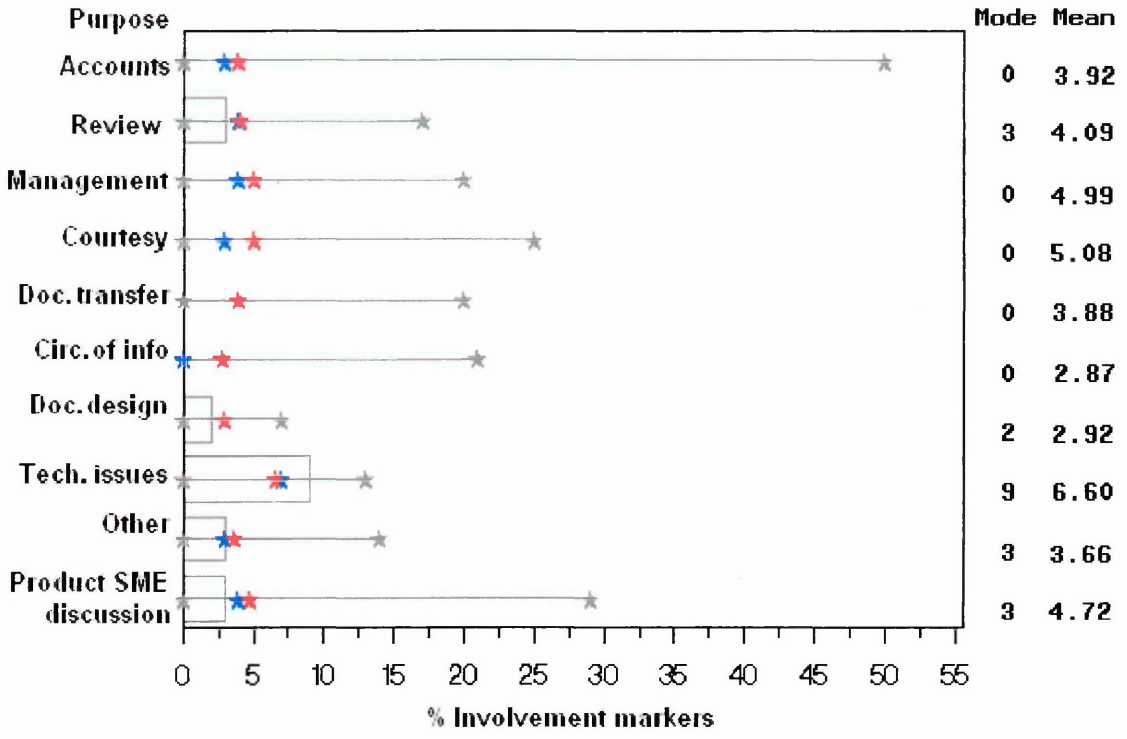

Figure 6-23: Differences by purpose for involvement

Legend: min ★— max —★ median ★ mean ★ mode □

Involvement by *Purpose* was high for *Management*, *Product* and *Review* discussions, and also for *Technology* and *Courtesy* issues. *Management*, *Product* and *Review* discussions accounted for 26%, 13% and 9% of the total communications respectively and were rated by the interviewee as of 'very high value' towards the goal of the team project (see appendix K, line 81-82). (*Product* and *Review* categories also had high elaborations in terms of body text word counts, accounted for by the complexity of the subject matter discussions.) *Courtesy* and *Technology* related issues accounted for only 4% and 1% of the total emails respectively and were both rated of 'very low value' towards the goal of the team project by the interviewee. In the commercial case study (chapter 5), *Courtesy* also showed relatively high involvement. *Courtesy* emails are essentially pro-social strategies and therefore have high levels of involvement, despite their perceived low value towards the goal. Most of the *Technology* emails in this study related to difficulties sending or opening documents and it is possible that feelings of frustration invoked higher involvement representations. Results from both this and the commercial case study suggest that the "Technology" category may have been better labelled "Difficulties". The high involvement for *Purposes* related to both the task (*Review* and *Product* discussions) and group maintenance (*Management* and *Courtesy*) demonstrates involvement in both task and social dimensions.

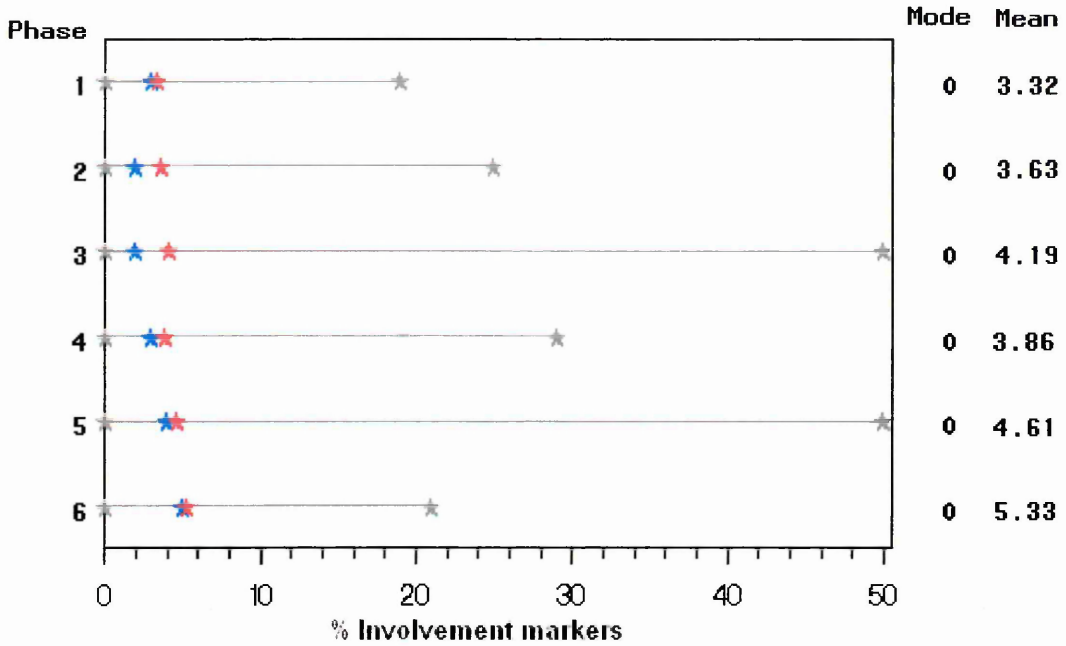

Figure 6-24: Differences by phase for involvement

Legend: min ★— max —★ median ★ mean ★ mode □

Figure 6-24 shows an overall trend of increasing involvement with socialisation phase. The socialisation phase is distinguished by increasing number of email communications, which were estimated on this project to account for 80% of the total communications (appendix K, line 67).

6.3.3 Solidarity and group cohesion

First person plural pronouns are interpreted as markers of solidarity or group cohesion. These markers showed variations with all the writing influences ($p = < 0.0001$). Frequency of these markers was relatively high, with more than one per email (total count 1072 in 866 emails). Additionally, the interviewee responded to questions on group belonging and cohesiveness in the post-analysis interview with the highest possible scores for all five questions, showing that her perception of group cohesion was high (appendix K line 111-113).

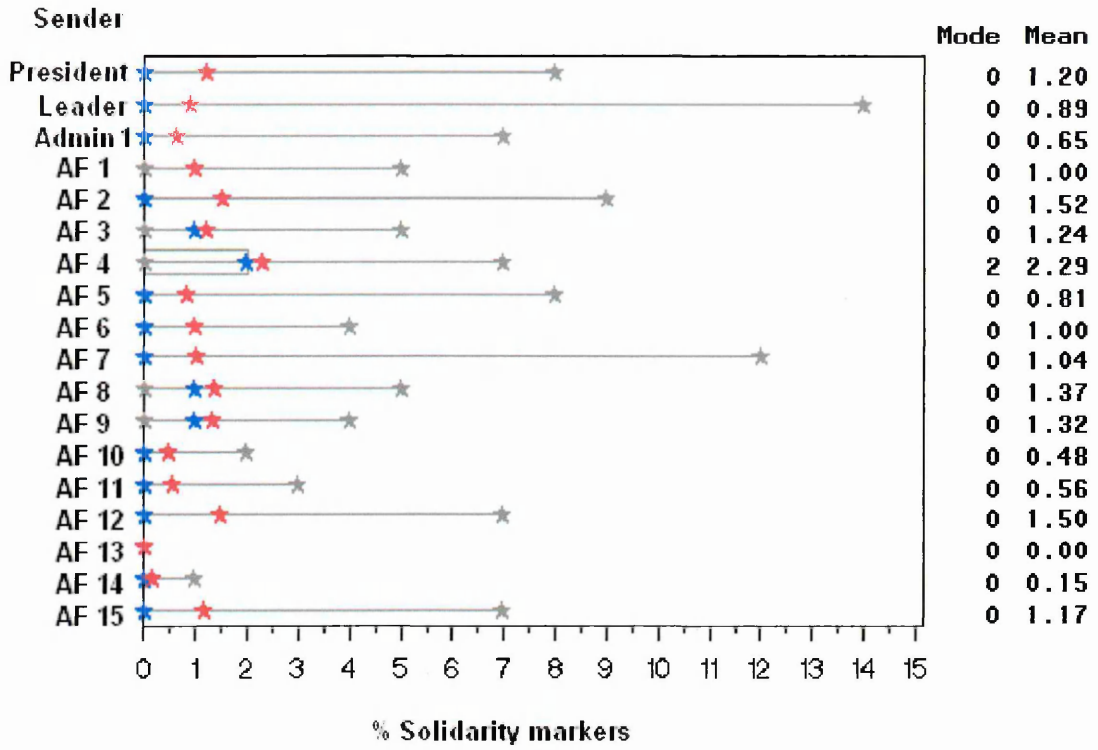

Figure 6-25: Differences by sender for solidarity

Legend: min max median mean mode

Solidarity markers (see Figure 6-25) were used most in emails by AF4, who the interviewee rated as being relatively less involved, less sociable, and more formal in the team (see Table 6-2, Table 6-3, Table 6-4). High expression of solidarity may be explained as a strategy related to purpose for this individual, who was responsible for the course organization in 2003, and who rated as having average socialisation in his functional role and being relatively new to the team (see appendix K line 45-46). Use of solidarity markers could therefore have been a pro-social strategy to encourage group cohesion and team cooperation, and also to create a sense of belonging to the group for this particular individual. It is also possible that such inconsistencies between interviewee perceptions of social behaviours and the email data may be introduced by the distortion of perceived FtF social behaviour as opposed to email communication behaviour. In teams which work mostly virtually, but have some FtF meetings, this factor confuses the use of team members' perceptions to interpret mediated behaviour patterns.

A further confusing effect is the choice to use first person singular or plural pronouns. AF13 showed the lowest frequency of solidarity markers, but also showed the highest frequency of involvement markers (see Figure 6-19). It is possible therefore that this *Sender's* individual style and preference for expression of involvement precludes her use of expressions of solidarity, explaining the two extremes in the solidarity and involvement representations.

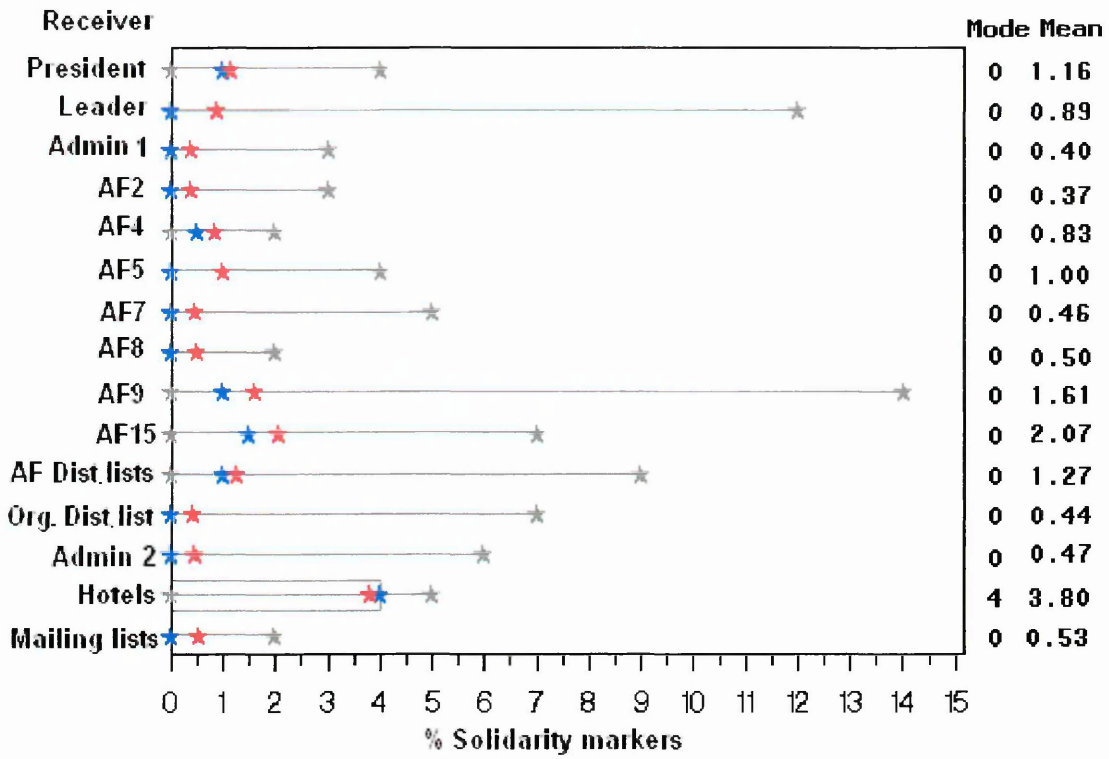

Figure 6-26: Differences by receiver for solidarity

Legend: min ★— max —★ median ★ mean ★ mode ▭
Admin = Conference administrator AF = Academic Faculty members
Leader = Course Leader; P = President; Org. Dist. List = Organizer's Distribution List

Writers adapted their expressions of solidarity according to the *Receiver* of the email, as shown in Figure 6-26. This example illustrates the caution required in interpreting communication behaviour markers in isolation. As representations of team cohesion, these markers might suggest expressions of solidarity with contacts who were least involved in the team project, i.e. hotel personnel and AF15. The interviewee rated AF15 as having relatively low involvement and sociability, and being a relatively formal team member (appendix K line 168-169). First person plural markers in this case, therefore, are more likely to reflect formal language strategies, where the plural voice represents an entity (such as the conference organizers or the Society) rather than an expression of solidarity between an individual (the writer) and other team members. Nevertheless, the markers still show a significant variation in style of communication with recipient of the email, showing adaptation by writers according to the *Receiver*.

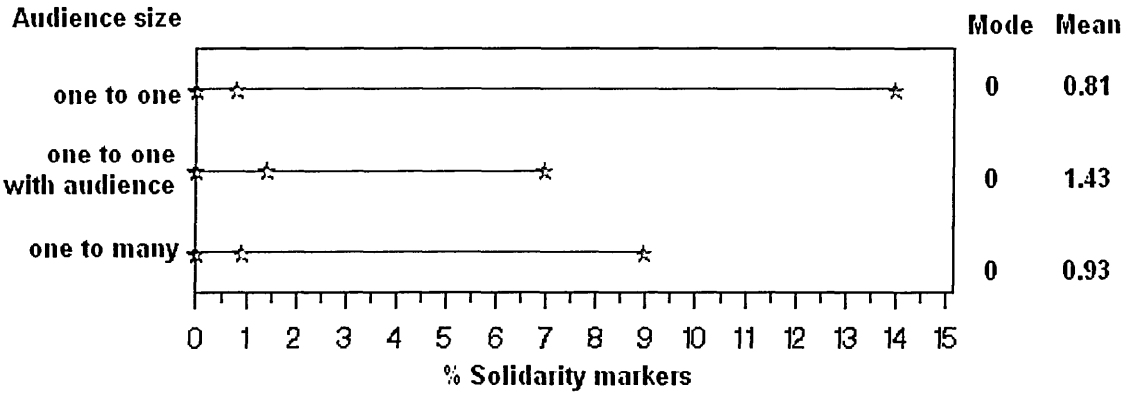

Figure 6-27: Differences by audience size for solidarity

Legend: min ☆— max—☆ median ★ mean ☆ mode ▭

From Figure 6-27 we can see that expressions of solidarity are greater with an audience than for one to one communications without an audience or for emails addressed to more than one individual. This trend was also observed in the Namahn data (see section 5.3.4). In the Society data, the trend may be explained by communications to mailing lists, i.e. to non-team members.

Solidarity by direction (Figure 6-28) is relatively high for communications between the President and Academic Faculty members, and amongst the Academic Faculty members, and relatively low in communications involving the Administrator. This interpretation of solidarity is reinforced by the interviewee's comment:

> *It's difficult with [Administrator], because she is like an outsider and was dealing with completely different issues. I don't think we were particularly formal with each other. It's difficult to categorize her. She wasn't actually part of the team doing the course work; she was there for more practical things (Society interviewee: appendix K line 175).*

The absence of solidarity markers in emails from the Course Leader to the Administrator may be explained by the higher FtF contact between these individuals. The Administrator was the conference organizer and FtF meetings were held on a monthly basis between these individuals (appendix K line 62). This may have altered the communication behaviour in emails in the sense that team building strategies, expressed through solidarity markers, were not necessary for communications between the Course Leader and Administrator; these individuals established their relationship in FtF meetings. On the other hand, the interviewee's comments above suggest that this individual did not have the solidarity experienced by the rest of the team members, the differentiation pertaining more to the functional roles (administrative rather than academic) than to the client-supplier relationship between the course organizer and the course designers.

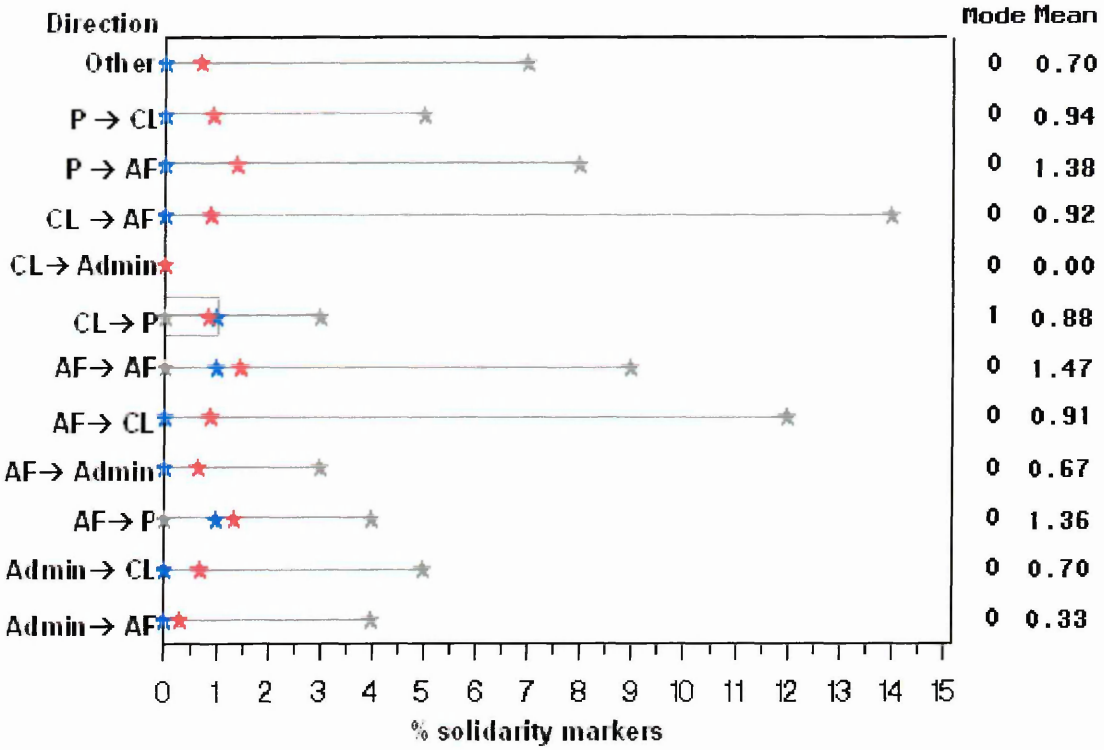

Figure 6-28: Differences by direction for solidarity

Legend: min ★— max —★ median ★ mean ★ mode ▭

Admin = Conference administrator AF = Academic Faculty members
CL = Course leader P = President

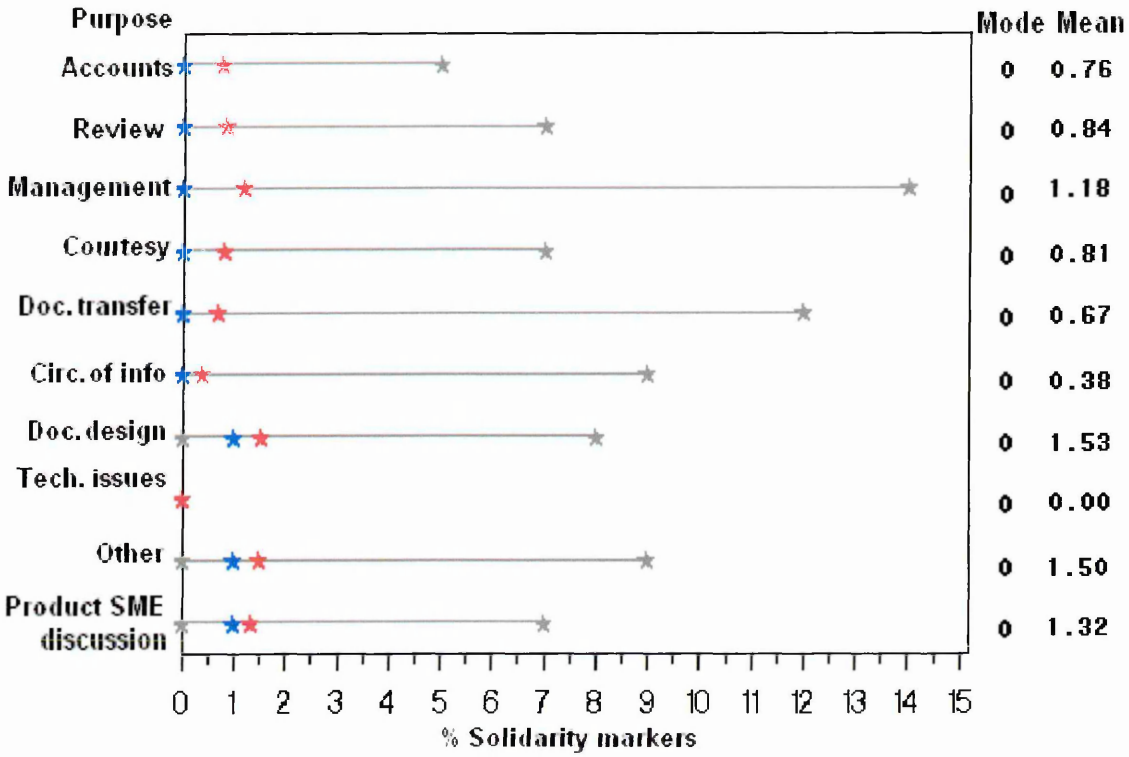

Figure 6-29: Differences by purpose for solidarity

Legend: min ★— max —★ median ★ mean ★ mode ▭

Representations of solidarity also varied with the *Purpose* of an email. Highest solidarity was expressed in emails over *Management*, *Other* Society business and *Product* and subject matter discussions, and also for *Document* Design issues, which included some discussion of potential publication channels. Thus solidarity was expressed for group maintenance (*Management*) and task related issues (e.g. *Product* discussion).

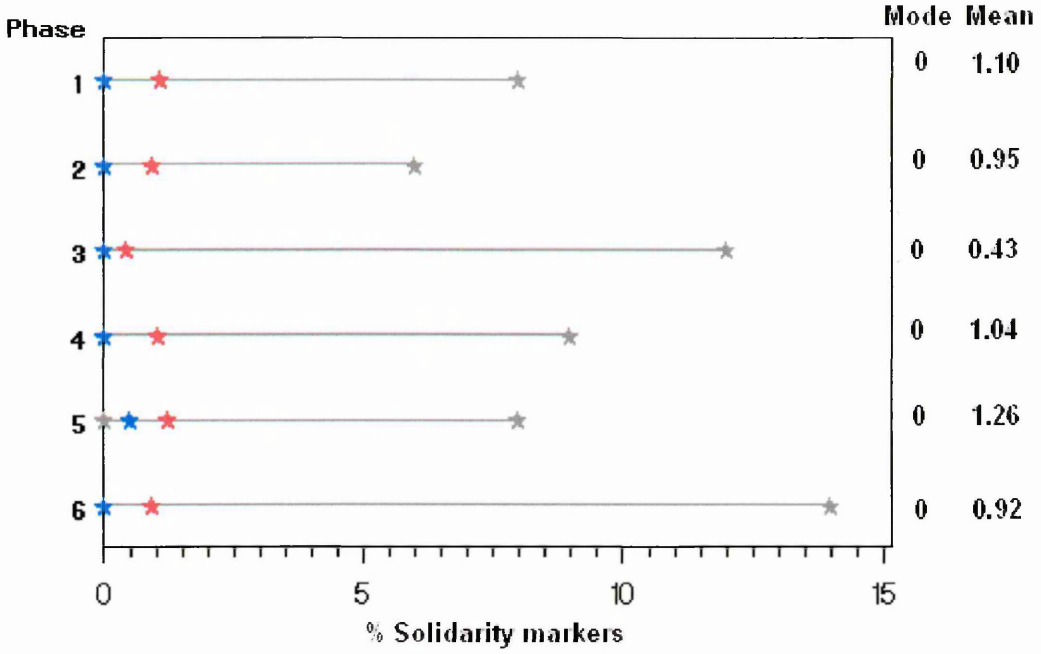

Figure 6-30: Differences by phase for solidarity

Legend: min ★ max ★ median ★ mean ★ mode

Similar to elaborations in body text, open and close greeting word counts (see Figure 6-6, Figure 6-17, Figure 6-18), solidarity appeared to increase after the 2002 conference, which happened towards the end of phase three. The interviewee commented on both task and social dimension changes after the conference, with task goals being re-evaluated and sociability increasing.

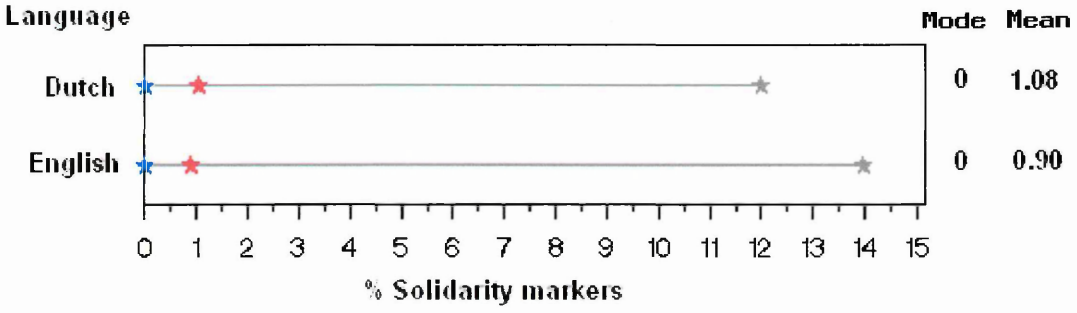

Figure 6-31: Differences by language for solidarity

Legend: min ★ max ★ median ★ mean ★ mode

Finally the Kruskal-Wallis result suggested that there was a difference between expressions of solidarity by *Language*. The means suggest a higher use in the Dutch than English emails. It is difficult to evaluate whether this reflects a stronger feeling of solidarity amongst Dutch writers, due to confusion of representation of this marker either as an expression of solidarity or representation of an official entity, such as 'we, the conference organizers'. Nickerson's (2000) research has shown a tendency for Dutch writers to use 'we' as a pronoun excluding the receiver (see section 2.3.6). In the Society data, there were 10 Dutch emails addressed to hotels, which used almost identical wording, for example:

> ***Wij*** *verwachten een versnelling in de reservaties aangezien* ***wij*** *deze dagen al* ***onze*** *sprekers aan het bevestigen zijn en ook alle congresgangers aan het aansporen zijn om zo snel mogelijk een hotelboeking te maken (Society project, email text reference 6648).*
>
> *English translation: We expect increasing reservations as we are currently confirming all our speakers and also encouraging all the conference attendees to book their hotel reservations as quickly as possible (Society project, email text reference 6648).*

This example may represent a sense of solidarity of the client (we, the conference organizers), in the potential client-supplier relationship (conference organizer-hotel) but does not share expressions of solidarity with the *Receiver(s).*

6.3.4 Formality score: greeting and signature style

The formality scale runs from a minimum score of 3 (=low formality) to a maximum of 9 (=high formality) comprising scores up to the value of 3 for each of the open and close greeting and manual signature. Formality varied between at least two of the categories for all writing influences ($p=<0.0001$ except for *Phase* $p= 0.0076$ and *Language* $p=0.0429$). I also asked the interviewee to rate each team member for formality and the results are shown in Table 6-3.

Table 6-3: Interviewee formality ratings, sorted by rating

Formality rating (1=most formal)	Team member
No rating	Administrator
No rating	Course Leader
1	AF1
1	AF5
1	AF6
1	AF13
2	President
2	AF8
3	AF4
4	AF3
4	AF12
5	AF15
6	AF11
7	AF10
8	AF14
9	AF9
10	AF2
11	AF7

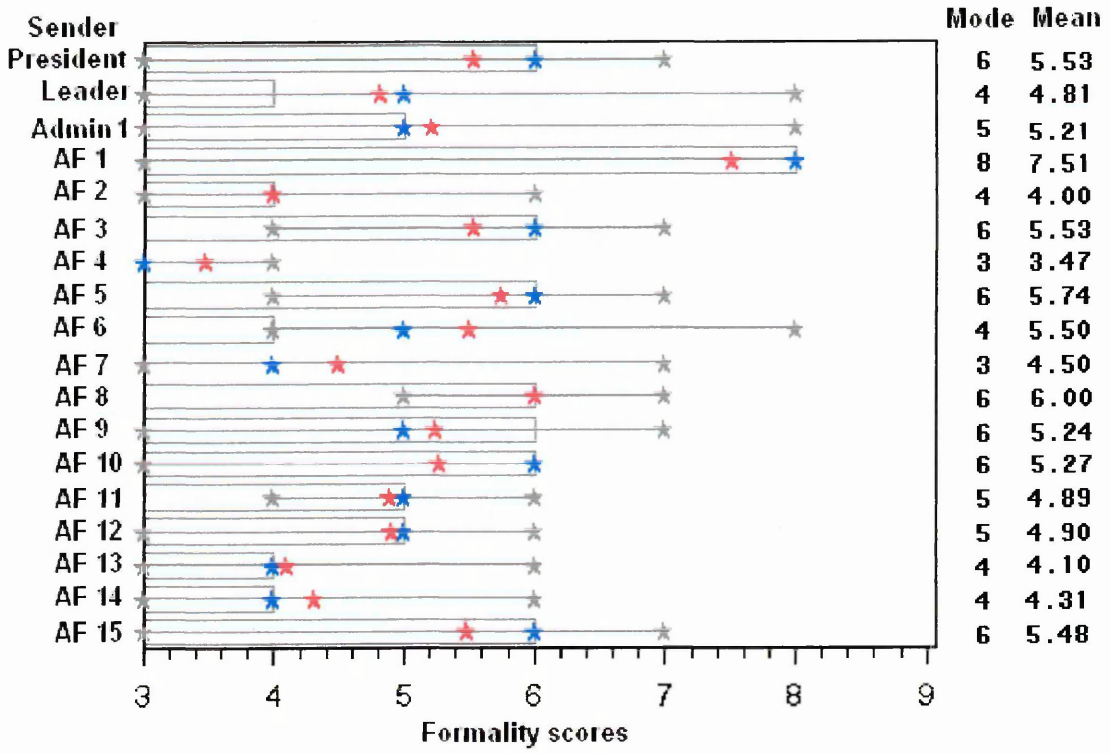

Figure 6-32: Differences by sender for formality scores

Legend: min max median mean mode

Figure 6-32 shows the variations in email formality style by *Sender*. The most formal writer of emails was AF1, who was also rated as highly formal by the interviewee. The least formal emails were sent by AF4. The interviewee rated AF4 as relatively formal, which does not fit with the interpretation from this formality scale. On the other hand the interviewee's perception of this individual as formal may explain the tendency of writers to use a formal style when writing to AF4 (see Figure 6-33).

Writers of emails adapted their formality style for *Receivers*, with the lowest formality (by modes) for AF2, AF7 and AF9. These individuals were also rated as least formal (see Table 6-2), so that writers adapted the formality style in emails to match the formality of the *Receiver*. The interviewee suggested that she expected less formality in emails from a colleague with whom she worked closely and had more FtF contact, rating this colleague as highly informal:

> *Regarding formality, [AF7] was my direct colleague (appendix K line 175).*

Thus style of email is also affected by the amount of FtF contact individuals have, which alters the socio-emotional component needed in emails to build and maintain relationships.

The highest formality by *Receiver* (by modes) was for the group of three individuals organizing the course and conference ("Org. Dist. List." in Figure 6-33), which corroborates the higher open greeting formality (Figure 6-12), lower involvement (Figure 6-21), relatively low solidarity (Figure 6-27) and higher formality (see Figure 6-34) of emails addressed to more than one person.

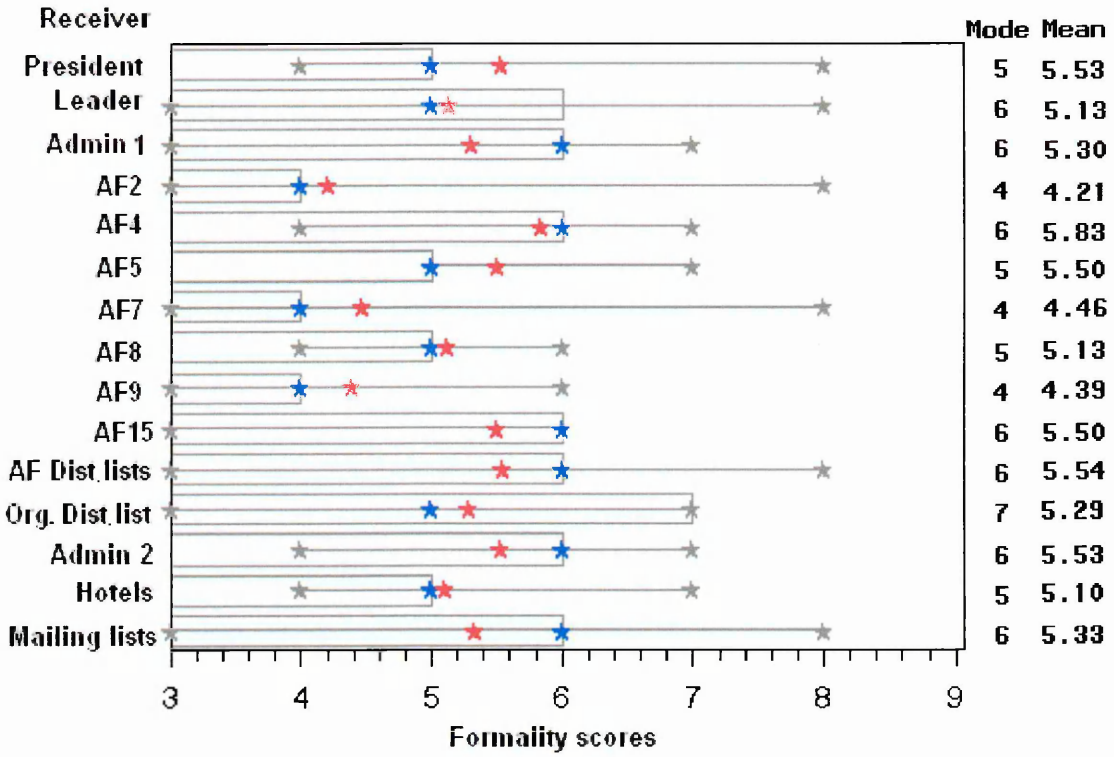

Figure 6-33: Differences by receiver for formality scores

Legend: min ★ max ★ median ★ mean ★ mode

Admin = Conference administrator AF = Academic Faculty members

Leader = Course Leader; P = President; Org. Dist. List = Organizer's Distribution List

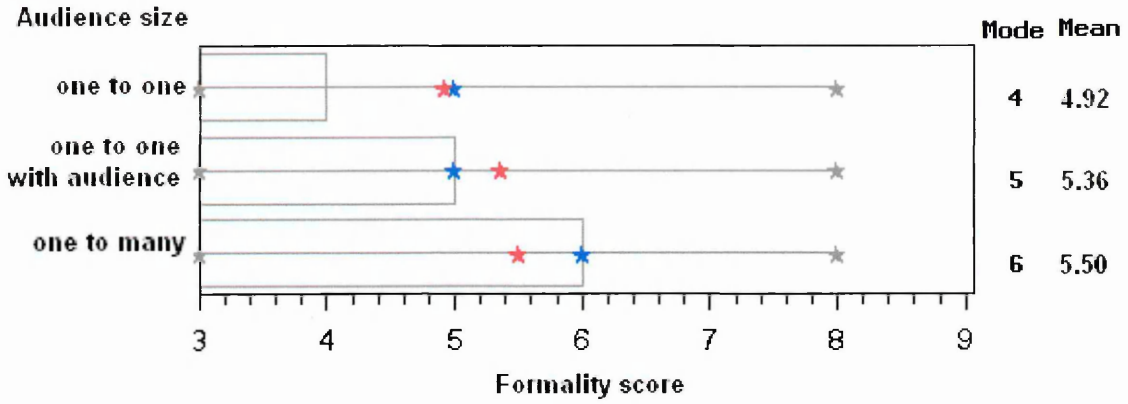

Figure 6-34: Differences by audience size for formality scores

Legend: min ★— max —★ median ★ mean ★ mode □

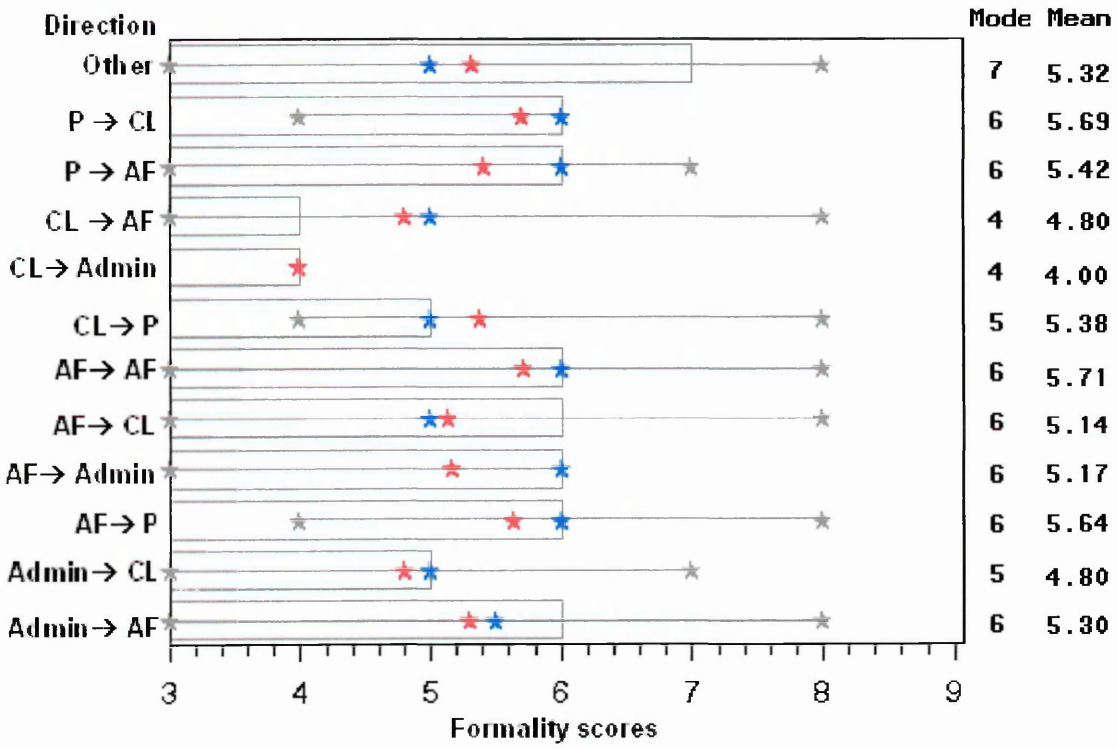

Figure 6-35: Differences by direction for formality scores

Legend: min ★— max —★ median ★ mean ★ mode □
Admin = Conference administrator AF = Academic Faculty members
CL = Course leader P = President

Formality varies with *Direction,* with the highest formality (by modes) in emails to non-team members, relatively high formality in emails to or from Academic Faculty members, and relatively low formality in emails involving the Course Leader and Administrator.

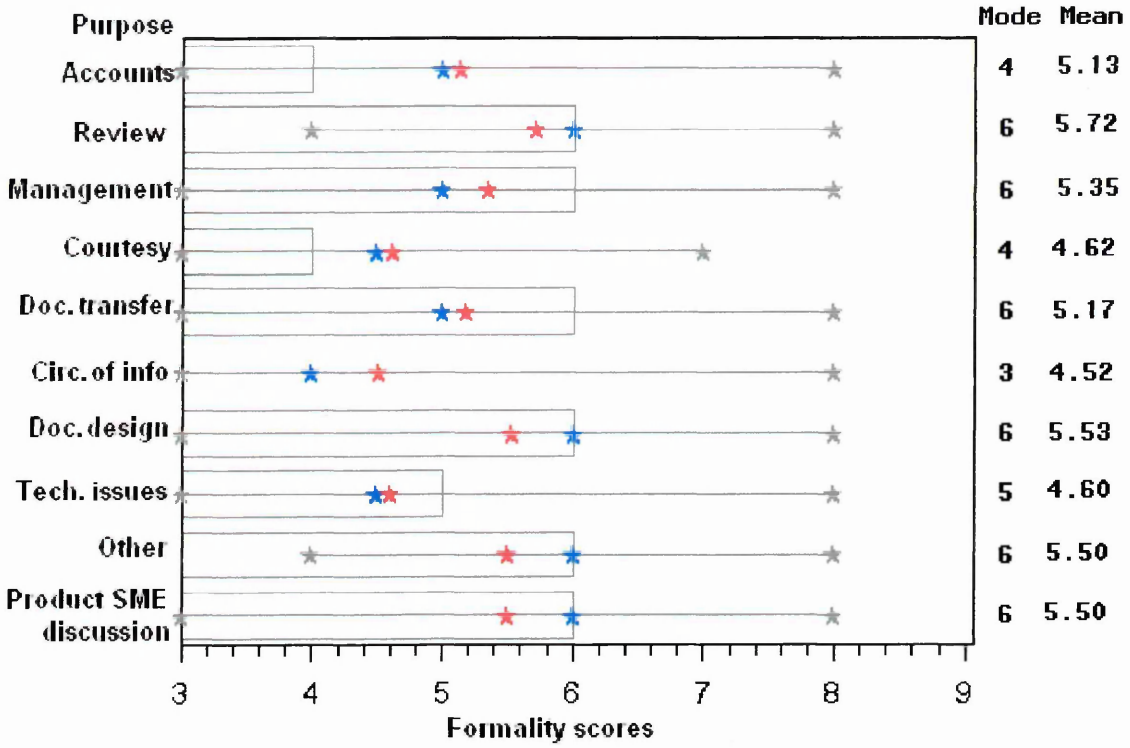

Figure 6-36: Differences by purpose for formality scores

Legend: min max median mean mode

Email formality varied by *Purpose (*Figure 6-36), with *Accounts, Courtesy* and *Circulation* of information having the lowest formality. *Review*, *Management*, *Document* transfer, *Document Design*, *Product* discussion and *Other* Society business all had relatively high formality.

Formality was most frequently lower in phases one and two, and higher for the next four phases.

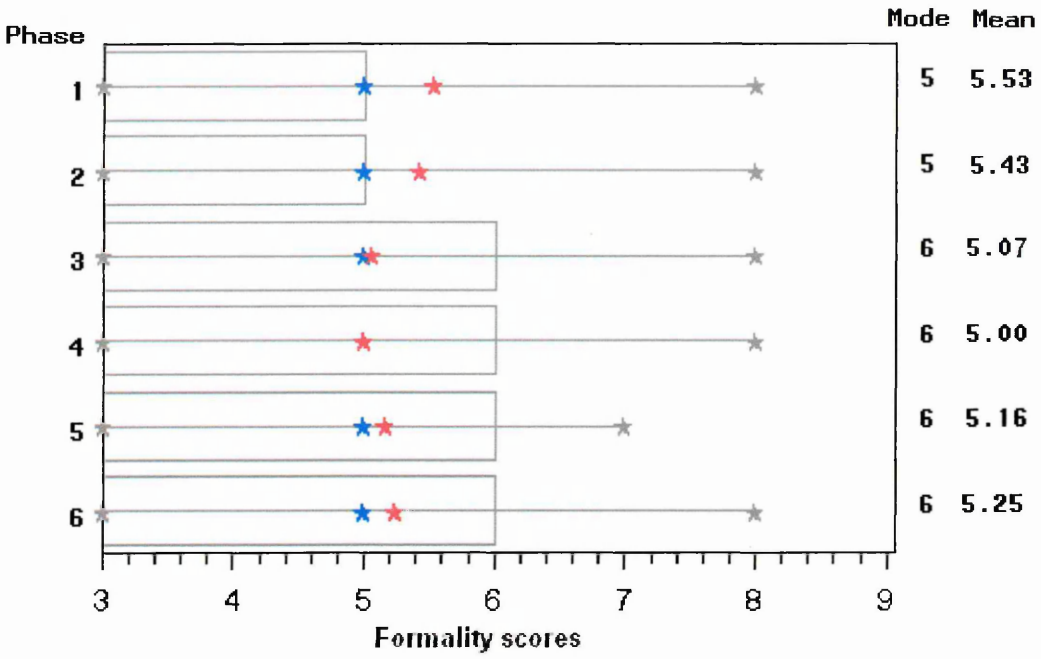

Figure 6-37: Differences by phase for formality scores

Legend: min ★— max —★ median ★ mean ★ mode □

Formality varied significantly with *Language*, with a lower mean score for English emails.

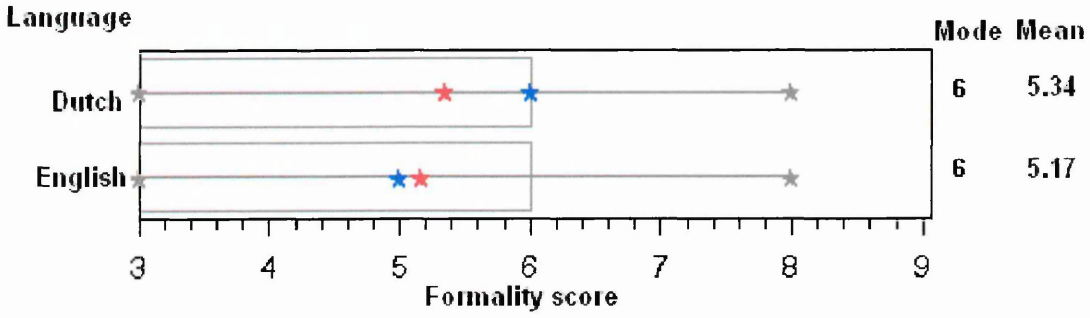

Figure 6-38: Differences by language for formality scores

Legend: min ★— max —★ median ★ mean ★ mode □

6.3.5 Sociability interpreted from social building units

Texts which might contribute towards the social development, interpersonal relationships and maintenance of the group were interpretively coded as "social building units". The total number of social building units per email varied between at least two categories for all the writing influences ($p=<0.0001$; for *Phase* $p= 0.0001$ and for *Language* $p= 0.0400$).

Interviewee ratings for sociability by team member are presented in Table 6-4.

Table 6-4: Interviewee ratings for sociability, sorted by rating

Sociable (1= most sociable)	Team member
No rating	Course Leader
No rating	Administrator
1	AF2
2	AF8
3	AF9
4	AF14
5	AF10
6	AF7
7	AF5
8	AF11
9	AF12
9	AF1
9	AF6
9	AF15
10	President
11	AF13
11	AF4
12	AF3

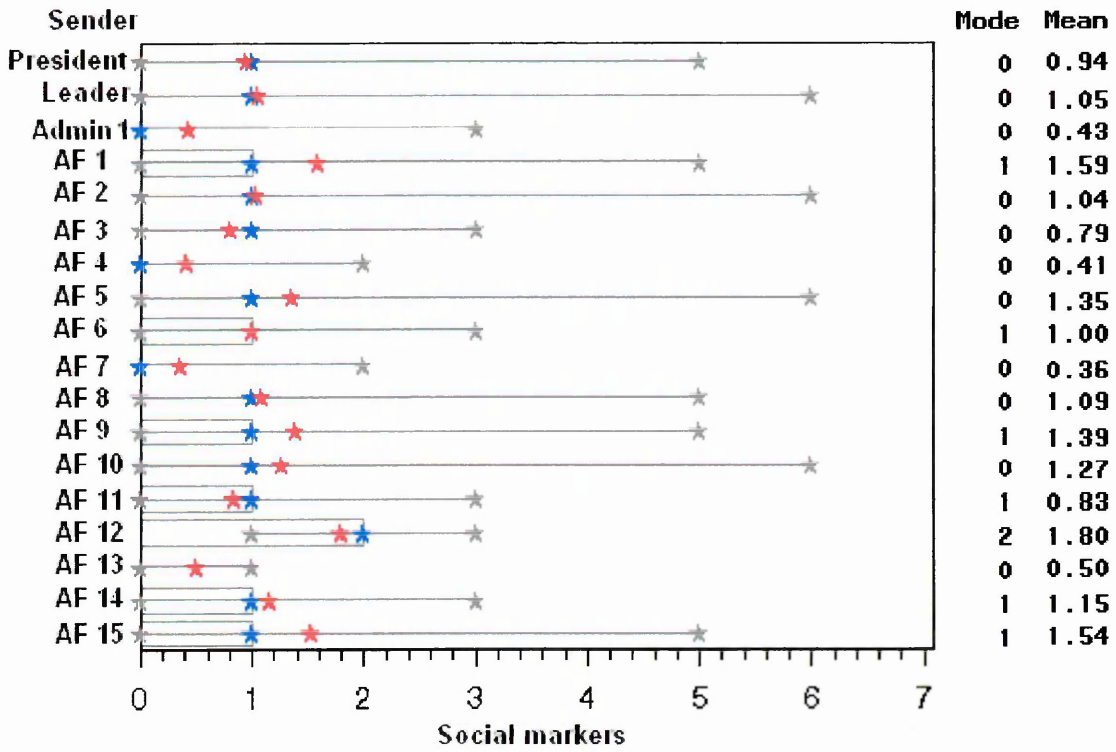

Figure 6-39: Differences by sender for social marker frequency

Legend: min ★ max ★ median ★ mean ★ mode

Expressions of sociability varied by *Sender;* AF12 showed the highest sociability (see Figure 6-39). *Senders* AF1, AF6, AF9, AF11, AF14 and AF15 all showed relatively high sociability.

However, the interviewee rated all of these *Senders* as relatively less sociable, except for AF9 and AF14. The most sociable by use of social markers in emails, AF12, wrote 10 emails and used 18 social markers. Out of these 18 markers, 16 were courtesy, 1 was a general social building unit and 1 apology. Thus AF12's high sociability is represented almost entirely by courtesy, which in this case may represent formality, and I give examples here:

> *Many thanks in advance (AF12 Text 2376).*
>
> *Thank you for your reply and for your participation on the reading committee (AF12 Text 2733).*
>
> *...and we thank you for your participation in [the Society] 2003 Meeting (AF12 Text 2728).*
>
> *...and thanks for your contribution to this [Society] Conference (AF12 Text 2727).*
>
> *Thank you for your email (AF12 Text 2722).*
>
> *Thank you for your mail and your very precise review (AF12 Text 2720).*
>
> *Would you please consider this correction and give me your opinion (AF12 Text 2714).*

Omission of the courtesy markers in these extracts from the emails would change the tone of the communication, for example if the *Sender* had written "I received your mail and your very precise review". The courtesy markers contribute towards building and maintaining good relationships, but the component they contribute is slightly different to that contributed, for example, by self disclosure. This difference argues the case for not combining such markers in a general social building representation.

When I asked the interviewee whether she found certain individuals more sociable in their communications, she responded "yes probably". I then asked what elements of their communications led her to interpret sociability and she responded:

> *Inclusion of personal things in emails, e.g. someone saying they were pregnant etc. (appendix K line 123).*

For the remaining *Senders* who appear to be highly sociable measured by email markers, but who were not rated as highly sociable by the interviewee, the distributions of marker types suggest that the sociability is mainly represented by the general social building units and courtesy markers. AF14 and AF9, who were rated as relatively sociable by the interviewee, both used expressions of self-disclosure and AF14 also used humour (see Figure 6-40). AF6 included one expression of self-disclosure referring to a family crisis which affected meeting attendance, i.e. this was an exception disclosure and perhaps not representative of this individual's usual style.

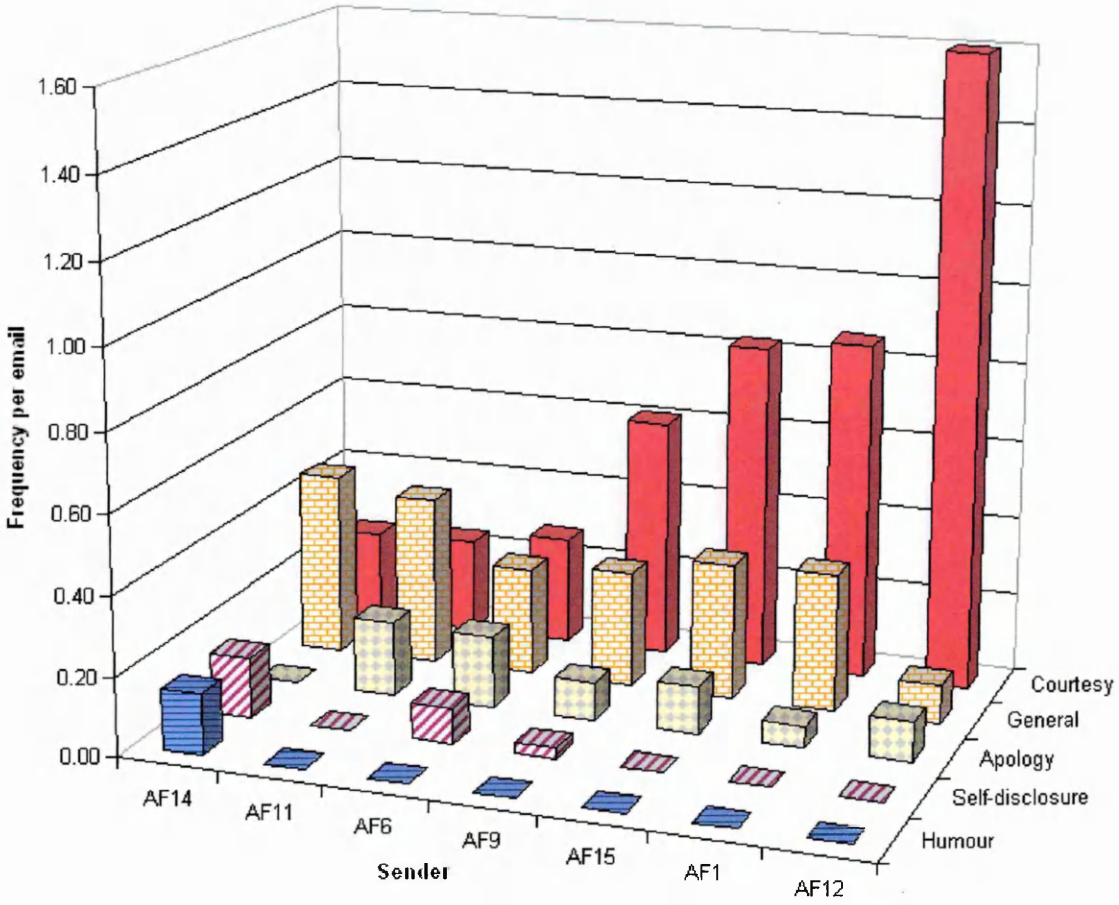

Figure 6-40: Breakdown of social marker types for senders with high sociability

Thus while all the social markers used relate to the attention paid by a *Sender* to the *Receiver* and the *Sender-Receiver* relationship, some have a more formal representation than others. These more formal markers still help to maintain good relations between team members, but may originate from socialised norms of behaviour and be adapted less for *Receivers*, than other social markers. For example, *Senders* may adapt their levels of self-disclosure for *Receivers*, but not their level of courtesy. The representation of these subcomponents of social markers requires more research to understand their roles either in conforming to norms or meeting the socio-emotional needs of communication.

Sender AF7 had the lowest sociability representation in emails, but was rated as relatively sociable by the interviewee. In a similar way to how the FtF meetings with the Administrator may have removed the need for team building strategy of expression of solidarity between the Course Leader and Administrator, FtF team working with the Course Leader, Administrator and AF7 may have removed the need for social building strategies in email communications amongst these three individuals.

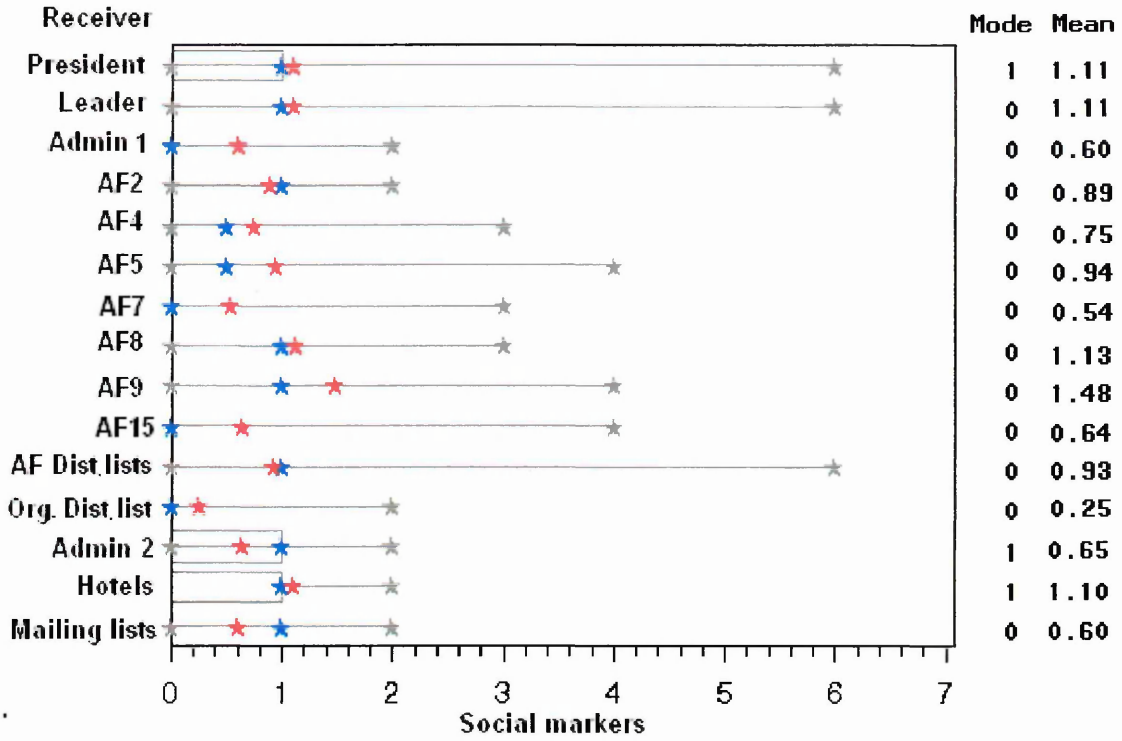

Figure 6-41: Differences by receiver for social marker frequency

Legend: min ★— max —★ median ★ mean ★ mode ▭
Admin = Conference administrator AF = Academic Faculty members
Leader = Course Leader; P = President; Org. Dist. List = Organizer's Distribution List

Writers adapted their communication styles for the *Receivers* of emails in terms of social building strategies. Marker frequency was highest in emails to AF9, who was rated by the interviewee as the third most sociable member of the team (see Table 6-4). This example suggests that writers attuned their use of social building markers according to knowledge of the sociability of the *Receiver*. The organizer's distribution list, which comprised three members with the most FtF contact, (Course Leader, Administrator and AF7) had the lowest need for social building units in email communications, and indeed received the lowest frequencies.

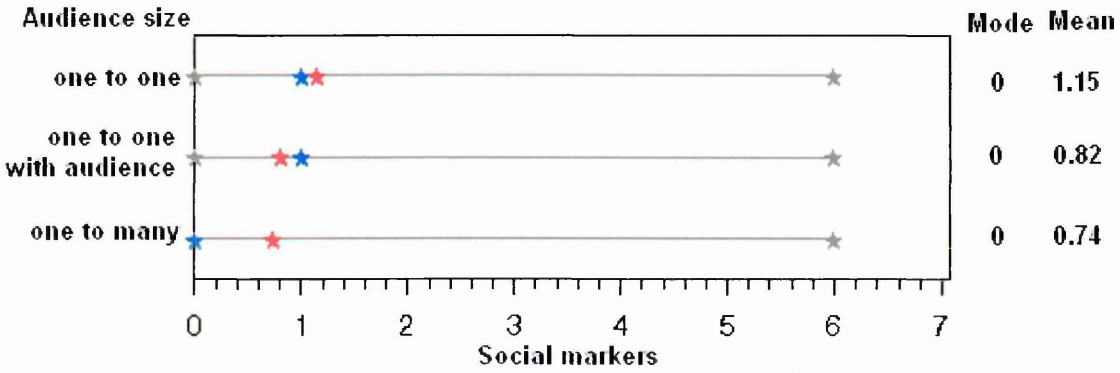

Figure 6-42: Differences by audience size for social marker frequency

Legend: min ★— max —★ median ★ mean ★ mode ▭

Social building markers decreased with increasing audience size, reflecting more use in private and interpersonal communications. The markers are therefore used more to build or maintain interpersonal relations than as solidarity building tools within the team.

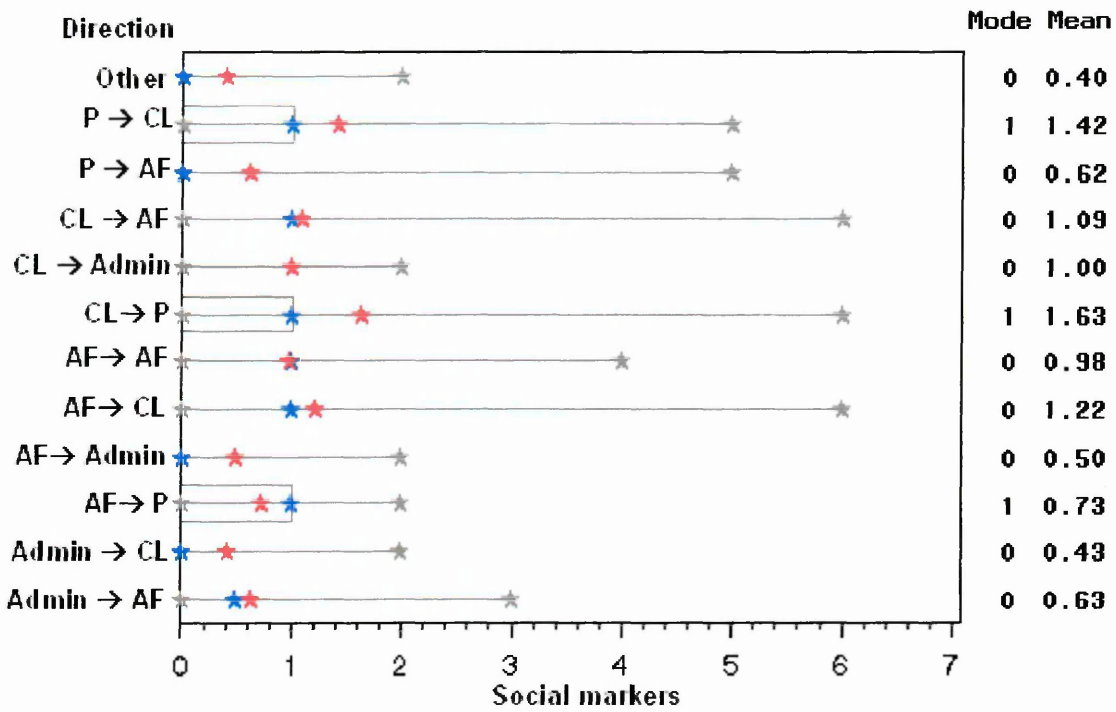

Figure 6-43: Differences by direction for social marker frequency

Legend: min ★— max —★ median ★ mean ★ mode ▭
Admin = Conference administrator AF = Academic Faculty members
CL = Course leader P = President

By direction (Figure 6-43), social markers were more frequent in emails from the Course Leader to the President and vice versa, and least frequent in the emails from the Academic Faculty to the Administrator. The interviewee commented on her relationship with the President as follows:

> *Things have changed a lot since 2002; this is difficult. Now I'm happy to face [the President] but at that time, he was like 'the big President of the Society' to me, you know (appendix K line 172).*

High social building markers in communications between the Course Leader and President therefore reflect social building strategies to develop the relationship between the two. Low social building marker frequency in communications between the Academic Faculty and the Administrator (who was the conference and course organizer) may reflect the client-supplier relationship or lack of knowledge of the *Receiver*, due to her being new to the team (appendix K, line 45-46).

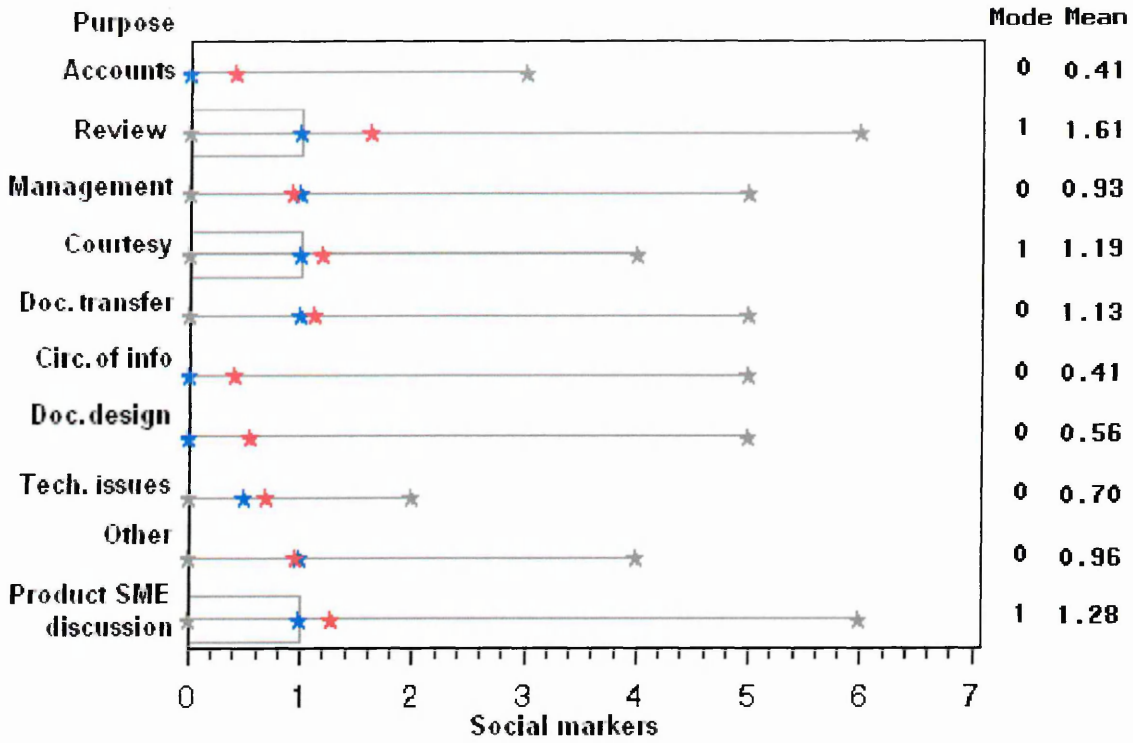

Figure 6-44: Differences by purpose for social marker frequency

Legend: min max median mean mode

Sociability varied with *Purpose* of an email as shown in Figure 6-44. *Review* and *Product* discussions, *Document transfer* and *Courtesy* had the highest sociability markers. Thus both task oriented purposes (e.g. *Review*) and a group maintenance oriented purpose (*Courtesy*) showed sociability. *Accounts* and *Circulation* of information showed the lowest sociability.

Sociability increased between phases 3 and 4 (Figure 6-45), which is attributable to occurrence of the 2002 conference at this time, after which the interviewee reported renewed task goal orientation (appendix K lines 72 and 155) and increased sociability:

> *Sociability increased after the [location] 2002 conference (Society interviewee: appendix K line 155).*

In general, there appears to be higher sociability in the second half of the project than the first half.

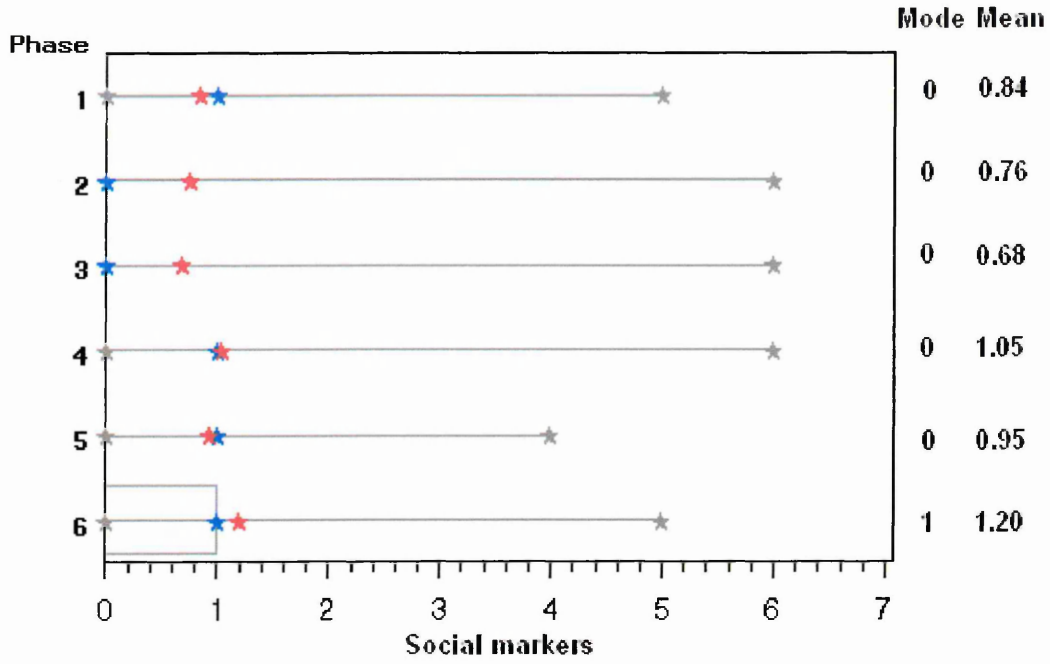

Figure 6-45: Differences by phase for social marker frequency

Legend: min ★ max ★ median ★ mean ★ mode

Finally, Dutch emails showed higher sociability than English emails (see Figure 6-46), which may be attributable to collocation of individuals. However, this is doubtful based on the arguments above that increased FtF contact appears to remove the need for social building strategies in email communications.

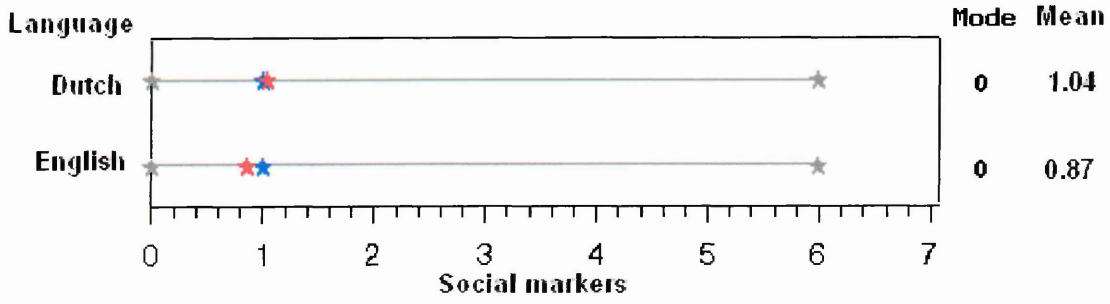

Figure 6-46: Differences by language for social marker frequency

Legend: min ★ max ★ median ★ mean ★ mode

6.4 Discussion

6.4.1 Overview

This case study aimed to identify the social dimension of a networked team writing project from content analysis of emails, addressing the following hypothesis:

H2 = Social dimensions of teams can be identified from email communications.

Email communications representing the discourse of a European Society developing a postgraduate clinical training course were analysed. Writing influences studied were *Sender, Receiver, Audience size, Direction, Purpose, Phase* and *Language*. Communication behaviour markers used to interpret the social dimension were also derived from the email content and included word count for body text, open and close greetings, % first person singular and plural pronouns, a formality score and frequency of social building units. These variables represented effort and value attributed to the communication, involvement, solidarity, formality and sociability. All the communication markers showed significant differences between at least two categories of the writing influence variables, with the following exceptions: there were no differences identified for word count or close greeting length with *Audience size*, and no differences identified for open greeting, close greeting or involvement with *Language*.

In the previous section, I analysed differences between categories of writing influences together with interviewee data to develop meaningful interpretations of some of the social dynamics represented by the communication markers. In this section, I draw together the results from different communication markers to build an overall profile of the balance between the task and social dimensions of the project. First I consider the results mapped against Nystrand's social interactive model of communication, and then discuss the evidence gathered pertaining to group cohesion. I then summarize the representations which emerged from the email data.

6.4.2 Email style and the social interactive model of writing

The aim of this research is to design an email analysis tool, which can be used to understand how team culture influences virtual team writing. Networked team members on writing projects need to communicate by email to achieve their team objective. Evaluating communication behaviour demonstrated in a team's written emails may therefore serve as a proxy means of predicting social interactivity in the writing of the final document. Knowing that the social dimension of team projects contributes positively to performance, this research focuses on the communication markers which represent both task and socio-emotional components of emails.

In Nystrand's social interactive model of writing, writers anticipate readers' needs, and meaning and interpretation is a shared social reality, the meeting of writer intentions and reader interpretations. In this study, writers adapted their communication behaviour for their intended *Receivers* using all the markers I studied, demonstrating the team's practice of social interactive writing behaviour.

Involvement may reflect one or both of the task and social dimensions in the communication. Involvement varied with both the *Receiver* and the *Purpose*, suggesting both task and socio-emotional involvement in this project.

Writers adapted their email style by *Direction* using all the markers, additionally demonstrating social interaction at an organizational level. This, together with the frequencies of greeting behaviour shown in Figure 3-9, Figure 3-10, and Figure 3-11 suggests the influence of social norms in addition to adaptation by *Sender* and *Receiver* and the *Sender-Receiver* relationship.

I conclude from the high number of marker adaptations with both *Purpose* and *Receiver*, and in particular the adaptation of involvement with both *Purpose* and *Receiver*, that the task and social dimensions of social interactive writing are evenly balanced in this project.

6.4.3 Group cohesion and sociability

First person plural pronouns are interpreted as markers of solidarity or group cohesion. These markers showed variations with all the writing influences ($p = < 0.0001$). Frequency of these markers was relatively high, with more than one per email (total count 1072 in 866 emails).

The interviewee responded with the highest possible scores for all five questions on team cohesion, showing that her perception of group cohesion was high. Additionally when responding to the question on whether she felt that she was really a part of the team, the interviewee commented:

> *That's the nice thing about this course (appendix K line 112-113).*

I conclude from both the email and the interviewee data that the group cohesion was high on this project. The high group cohesion and the fact that all the communication markers were adapted both for *Purpose* and *Receiver*, suggests an even balance of the task and social dimensions of this project. This profile of an even social to task balance in this project will be compared with a commercial project in chapter 7.

6.4.4 Other social dimension interpretations

Senders

Effort and value assigned to a communication in terms of elaboration was partly individualistic, but independent of the overall activity contributed by an individual to the project. Individuals vary in the style which they most frequently use for both open and close greetings. However, *Senders* who use shorter open greetings do not necessarily also use a shorter close greeting. Thus, styles of open and close greeting vary by individual, and individuals may have different open and close greeting styles, so that interpretations for overall email style cannot be based on one or other greeting alone.

Involvement, solidarity, formality and sociability all varied with *Sender.* Interviewee ratings corroborated email data and also highlighted some complexities in dual marker interpretations. Involvement represents both task and social involvement, solidarity and involvement markers are mutually exclusive, and sociability is comprised of multiple markers representing slightly different socio-emotional strategies.

Audience size

Involvement and sociability decrease and formality increases with *Audience* size, reflecting the closeness of one to one interpersonal communications (see Figure 6-47). Both formality and open greeting length increased with audience size, reinforcing the interpretation of increased

formality with increasing audience size. Solidarity was higher with an audience than for one to one communications without an audience, but emails addressed to several individuals showed lower solidarity, reflecting communications to mailing lists, i.e. to non-team members.

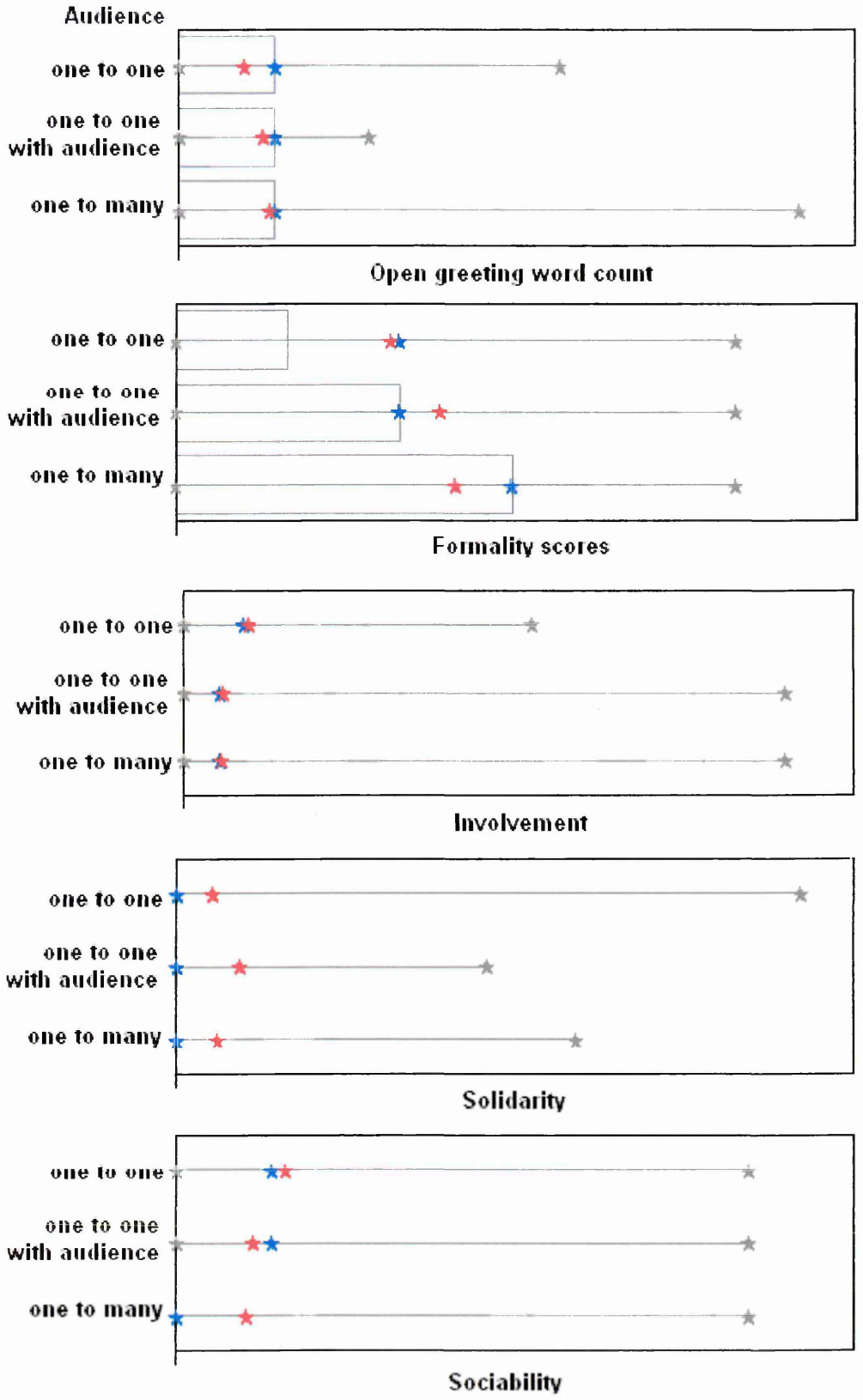

Figure 6-47: Communication behaviour by audience size

Direction

There were higher elaborations between the Course Leader and President and between members of the Academic Faculty, and lower elaborations between the Course Leader and Administrator, and vice versa. This reflects the high need for elaboration of complex subject matter content, and lower need for elaboration over less complex administrative and organizational issues (also substantiated by the elaboration profile by *Purpose)*.

Emails from the Course Leader to the President and vice versa, and from the Course Leader to the Administrator had the highest involvement by *Direction*.

Solidarity by *Direction* was relatively high for communications between the President and Academic Faculty members, and amongst the Academic Faculty members, and relatively low in communications involving the Administrator. The absence of solidarity markers in emails from the Course Leader to the Administrator is explained by the higher FtF contact between these individuals. This may have altered the communication behaviour in emails in the sense that team building strategies, expressed through solidarity markers, were not necessary for communications between the Course Leader and Administrator; these individuals established their relationship in FtF meetings. However, the interviewee described the Administrator as external to the team, which accurately reflects organizational differences, despite the fact that the Administrator was the most active communicator on the project.

Formality varies with *Direction,* with the highest formality in emails to non-team members, relatively high formality in emails to or from Academic Faculty members, and relatively low formality in emails involving the Course Leader and Administrator.

Social markers were more frequent in emails from the Course Leader to the President and vice versa, and least frequent in the emails from the Academic Faculty to the Administrator. Interviewee feedback on this relationship suggests that high social building markers in communications between the Course Leader and President reflected social building strategies to develop the relationship. Low social building marker frequency in communications between the Academic Faculty and the Administrator may reflect the client-supplier relationship or lack of knowledge of the *Receiver* due to the Administrator being new to the project.

Thus in conclusion, marker adaptations by *Direction* tend to profile two factors, the content of emails (administrative versus subject matter expertise) and the client-supplier or distinctness of the two organizations, the Society and the Conference Organizers. This brings me back to the difficulty of defining virtual discourse communities (discussed in section 2.2.7). The Administrator was the most active communicator in this set of emails, and was involved in many purposes related to document creation, but was not considered to be a member of the team by the interviewee.

Purpose

Elaboration is higher for *Purposes* involving discussion over the Society's subject matter expertise, and lower when emails are used as transfer agents or for group maintenance issues such as *Management* and *Courtesy* emails.

Open greetings were also longest for *Review* and *Product* discussions, and lowest for *Courtesy* messages and *Circulation* of information (see Figure 6-15). Close greetings showed a similar trend with long close greetings for *Review* and *Product* discussions (see Figure 6-16) and short close greetings for *Circulation* of information.

Involvement was high for *Management*, *Product* and *Review* discussions, and also for *Technology* and *Courtesy* issues. The high involvement for *Purposes* related to both the task (*Review* and *Product* discussions) and group maintenance (*Management* and *Courtesy*) dimensions demonstrates involvement in both dimensions.

Highest solidarity was expressed in emails over *Management*, *Other* Society business and *Product* and subject matter discussions, and also for *Document* Design issues, which included some discussion of potential publication channels. Thus solidarity was expressed for group maintenance (*Management*) and task related issues (e.g. *Product* discussion).

Accounts, Courtesy and *Circulation* of information showed the lowest formality. *Review*, *Management*, *Document* transfer, *Document Design*, *Product* discussion and *Other* Society business emails were relatively formal.

Review and *Product* discussions, *Document transfer* and *Courtesy* had the highest sociability markers. Thus both task oriented purposes (e.g. *Review*) and a group maintenance oriented purpose (*Courtesy*) showed sociability.

Thus to conclude on adaptation of communication markers by *Purpose*, *Courtesy* and *Management*, socially oriented *Purposes* showed lower elaboration, and Courtesy also showed lower formality. Involvement, solidarity and sociability scores were high in *Purposes* on both the task and social dimensions.

Phase

The interviewee reported renewed task goal orientation and increased sociability (appendix K lines 72 and 155) after the 2002 conference at a point in time between socialisation phase three and four. Phase four showed an increase in effort and value attributed to communications (in both body texts and greetings), and increases in solidarity and sociability. Thus FtF meeting, and accomplishment of the course and conference in 2002 improved the task orientation and sociability of the project and these improvements were visible from the communication markers. Additionally, involvement showed an overall increase with socialisation phase in this project.

6.4.5 Complex marker representations

In this study the interviewee was asked to rate team members on relative activity, involvement, sociability and formality. Many of the interviewee ratings corroborated findings from the email data, and also helped to identify complexities in the markers, which I discuss here.

In this data, open and close greeting behaviour is not adapted consistently. Variations in style of open and close greeting thus suggest that the markers differ in their socio-emotional components. Inconsistencies between interviewee ratings and formality interpreted from greeting lengths also highlight the dual representation of greeting lengths as markers of formality and markers of effort and value attributed to the *Sender-Receiver* relationship. Short open greeting styles may represent relatively low formality or relatively low effort and value attributed to the communication or to the relationship with the *Receiver*. Additionally, greeting length varies not only with *Sender* and *Receiver*, but also with influences on behaviour such as *Direction* and *Audience* size, showing that greetings are adapted to expected norms of behaviour.

Social building units are also complex and highly informative markers. An individual (AF12) rated by the interviewee with low sociability and involvement, but high formality had the highest expression of sociability in his emails, represented mainly by courtesy markers. High frequencies of general social building markers and courtesy markers by individuals had not encouraged the interviewee to rate individuals as particularly sociable. The interviewee felt that people who included personal information about themselves were more sociable. Indeed, the two individuals who she had rated as sociable, and who did have high frequencies of social markers, had also included self disclosure markers in their emails. Different types of social building marker thus represent different strategies in pro-social behaviour, some being influenced more by individuality and the *Sender-Receiver* relationship, and others being influenced by social norms of behaviour.

Social building units also have dual representations in being representative of the status of a relationship and the strategy used to build a relationship. It is possible that the different types of social building units used in this study may distinguish between existing sociability in interpersonal relationships and strategies to build relationships. For example, self disclosure may be more common in existing relationships, and courtesy and apology may be more common as social building strategies to develop relationships. The varying representations of different types of social building markers in relational communication require further research, which I discuss in section 6.5 and chapter 8.

Confusion in interpreting solidarity markers was introduced by the mutual exclusivity of involvement and solidarity markers. Sender AF13 had the highest involvement but lowest solidarity, which may be explained by her preference to use singular over plural pronouns, rather than being representative of her solidarity within the group. Additionally, first person plural markers can reflect formal language strategies, where the plural voice represents an entity

(such as the conference organizers or the Society) rather than an expression of solidarity between an individual (the writer) and other team members.

Finally, relative activity, task and socio-emotional involvement are all slightly different descriptors of an individual's behaviour. The involvement marker represents both task and socio-emotional involvement in an individual's contribution, but not the extent of the contribution in the project. Thus an individual can deliver only 1% of the project activities (measured by email frequency), but still show relatively high involvement from the email content.

These complexities of the markers and their interpretations mean that they cannot be interpreted in isolation or in a purely quantitative way. Rather they need to be searched in a holistic way, because each marker comprises slightly different and sometimes multiple representations, which contribute to the whole picture of an individual's style of email communication behaviour. In particular further research is required to find ways of systematically distinguishing between the underpinning pro-social strategies in uses of the social building and solidarity markers used in this research.

6.4.6 Face to face confounding effects

There was some evidence in this data that mediated communication behaviour may be influenced by the amount of FtF contact between individuals; writers with more FtF contact used solidarity and social building markers less in their email communications, suggesting that FtF meetings may remove the need for such strategies.

The absence of solidarity markers in emails from the Course Leader to the Administrator is explained by the higher FtF contact between these individuals. The Administrator was the conference organizer and FtF meetings were held on a monthly basis between these individuals (appendix K line 62). This may have altered the communication behaviour in emails in the sense that team building strategies, expressed through solidarity markers, were not necessary for communications between the Course Leader and Administrator; these individuals established their relationship in FtF meetings.

AF7 had the lowest sociability representation in emails, but was rated as one of the most sociable by the interviewee. In a similar way to how the FtF meetings with the Administrator may have removed the need for team building strategy of expression of solidarity between the Course Leader and Administrator, FtF team working with the Course Leader and AF7 may have removed the need for social building strategies in email communications and thus influenced the representation of sociability through markers in the emails. The interviewee also suggested that she expected less formality in emails from the colleague (AF7) with whom she worked closely and had more FtF contact, rating this colleague as highly informal.

The organizer's distribution list, which comprised three members with the most FtF contact, (Course Leader, Administrator and AF7) had the lowest need for social building units in email communications, and indeed received the lowest frequencies.

Additionally the interviewee's perceptions may have been coloured by the relative FtF contact she had with different individuals. She rated AF2 as the fourth most active on the project, although this individual was only represented by 2% of the emails. When shown the email frequency chart, and asked whether she could explain the discrepancy between perceived activity and email activity, she commented:

> *We must have discussed more by phone. I travelled to meet her once a year and she came here once a year – probably that was the person I had more face to face communications with. Out of the 5% face to face, she probably accounted for 90% (appendix K line 200).*

Thus style of email is also affected by the amount of FtF contact individuals have, which alters the socio-emotional component needed in emails to build and maintain relationships. Additionally when using perception data to research mediated communications, we need to take into consideration not only the relative amount of mediated communication on the project, but also the relative amount of mediated communication per team member.

6.4.7 Culture and language

Emails written in the Dutch language had higher elaboration, solidarity and sociability.

Whether the solidarity markers reflected a stronger feeling of solidarity amongst Dutch writers is uncertain, due to confusion of representation of this marker either as an expression of solidarity or representation of an official entity, such as 'we, the conference organizers'. The higher sociability in Dutch emails may be attributable to collocation of individuals. However, this is doubtful based on the arguments above that increased FtF contact appears to remove the need for social building strategies in email communications.

Formality varied significantly with *Language*, with a lower mean score for English emails.

6.5 Chapter review

This study focused on the hypothesis:

> *H2 = Social dimensions of teams can be identified from email communications.*

Email communications representing the discourse of a European Society developing a postgraduate clinical training course were analysed. Writing influences studied were *Sender, Receiver, Audience size, Direction, Purpose, Phase* and *Language*. Communication behaviour markers used to interpret the social dimension were also derived from the email content and included word count for body text, open and close greetings, % first person singular and plural pronouns, a formality score and frequency of social building units. These variables represented effort and value attributed to the communication, involvement, solidarity, formality and sociability. All the communication markers showed significant differences between at least two categories of the writing influence variables, with the following exceptions: there were no differences identified for word count or close greeting length with *Audience size*, and no differences identified for open greeting, close greeting or involvement with *Language*.

In Nystrand's social interactive model of writing, writers anticipate readers' needs, and meaning and interpretation is a shared social reality, the meeting of writer intentions and reader interpretations. In this study, writers adapted their communication behaviour for their intended *Receivers* using all the markers I studied, demonstrating the team members' practice of social interactive writing behaviour.

Involvement may reflect one or both of the task and social dimensions in the communication. Involvement varied with both the *Receiver* and the *Purpose*, suggesting both task and socio-emotional involvement in this project.

Writers adapted their email style by *Direction* using all the markers, additionally demonstrating social interaction at an organizational level. This, together with the frequencies of greeting behaviour shown in Figure 3-9, Figure 3-10, and Figure 3-11 suggests the influence of social norms in addition to adaptation by *Sender* and *Receiver* and the *Sender-Receiver* relationship.

Solidarity shown in communications varied in this project with all the writing influences, was relatively high in terms of marker frequency, and was perceived to be high by the interviewee.

I conclude from the high solidarity, high number of marker adaptations with both *Purpose* and *Receiver*, and in particular the adaptation of involvement with both *Purpose* and *Receiver*, that the task and social dimensions of social interactive writing are evenly balanced in this project. This profile of an even social to task balance in this project will be compared with a commercial project in chapter 7.

At a micro level, I analysed the communication behaviours against the writing influences studied to interpret trends in social behaviour using feedback from the interviewee. Social aspects of the project which emerged from the profiles are:

- Communication behaviour by *Direction* profiled two factors, administrative versus subject matter expertise functions of emails, and the fact that there were two separate organizations involved, the Society and the Conference Organizers. This highlighted the difficulty of defining a virtual discourse community. The Administrator was the most active communicator in this set of emails, and was involved in purposes related to document creation, but was not considered by the interviewee to be part of the team.
- *Courtesy* and *Management*, which are socially oriented *Purposes* showed lower elaboration, and *Courtesy* also showed lower formality. Involvement, solidarity and sociability scores were high in *Purposes* on both the task and social dimensions.
- The interviewee reported renewed task goal orientation and increased sociability after the 2002 conference at a point in time between socialisation phase three and four. Phase four showed an increase in effort and value attributed to communications, and an increase in solidarity and sociability. Thus FtF meeting, accomplishment of the course and conference in 2002 improved the task orientation and sociability of the project and these improvements were visible from the communication markers.

- Involvement showed an overall increase with socialisation phase in this project.

Complexities which need to be taken into account when interpreting communication behaviour and which emerged from analysis of interviewee feedback and the email data are:

- Open and close greeting behaviour does not adapt consistently; greeting length may represent both formality and value and effort attributed to communications, and varies with influences on behaviour such as *Direction* and *Audience* size, suggesting the influence of social norms as well as the Sender-Receiver relationship. Variations in style of open and close greeting suggest that the markers comprise different socio-emotional components.
- In depth analysis of the individuals with high frequencies of social markers suggests that the different types of marker may represent slightly different strategies in pro-social behaviour. For example, self-disclosure may be influenced more by individuality and the *Sender-Receiver* relationship, whereas courtesy may be influenced more by social norms of behaviour.
- Social building markers have dual representations in being representative of the status of a relationship and the strategy used to build a relationship.
- Solidarity and involvement marker representations may be confused by the possibility of only including one or the other, and solidarity markers may be used as formal representations of an entity rather than a representation of solidarity.
- Involvement markers represent both task and socio-emotional involvement in an individual's contribution, but not the extent of the contribution in the project.
- Email style is also affected by the amount of FtF contact individuals have, which alters the socio-emotional component needed in emails to build and maintain relationships, so that the relative amount of mediated communication per team member also needs to be taken into account in researching mediated communications.

These complexities of the markers and their interpretations mean that they cannot be interpreted in isolation, or in a purely quantitative way. Rather they need to be searched in a holistic way, because each marker comprises slightly different and sometimes multiple representations, which contribute to the whole picture of an individual's style of email communication behaviour. In particular further research is required to find ways of systematically distinguishing between the underpinning pro-social strategies in uses of the social building and solidarity markers used in this research.

The results of this case study show that the methodology of email content analysis delivers variables which clearly represent writing influences and communication behaviours, and that communication behaviours can be shown empirically to vary with the writing influences as predicted by the social interactive theory of writing. Interpreting the differences in

communication behaviours helped to develop meaningful interpretations of social behaviour on the project and to identify complexities in dual representations of communication markers.

Analysis at a micro level provides a useful problem solving tool in professional team projects. However, the main value of the tool is in comparing different projects at a macro level, taking adaptations of, and overall scores for formality, sociability, involvement and solidarity markers into account. These communication behaviours, representing social behaviour, do vary by *Sender*, *Receiver* and other writing influences, conforming to Nystrand's social interactive model of communication. This substantiates the study of these markers at a macro or group level to compare the social dimensions of projects. The profile of social behaviour derived from relational content and adaptations in relational metadiscourse, allows us to compare team writing projects and explore team behaviours, which may influence virtual team writing. This concept is explored in the next chapter.

7. H3: Comparing academic and commercial contexts – towards a causal model

7.1 Research focus

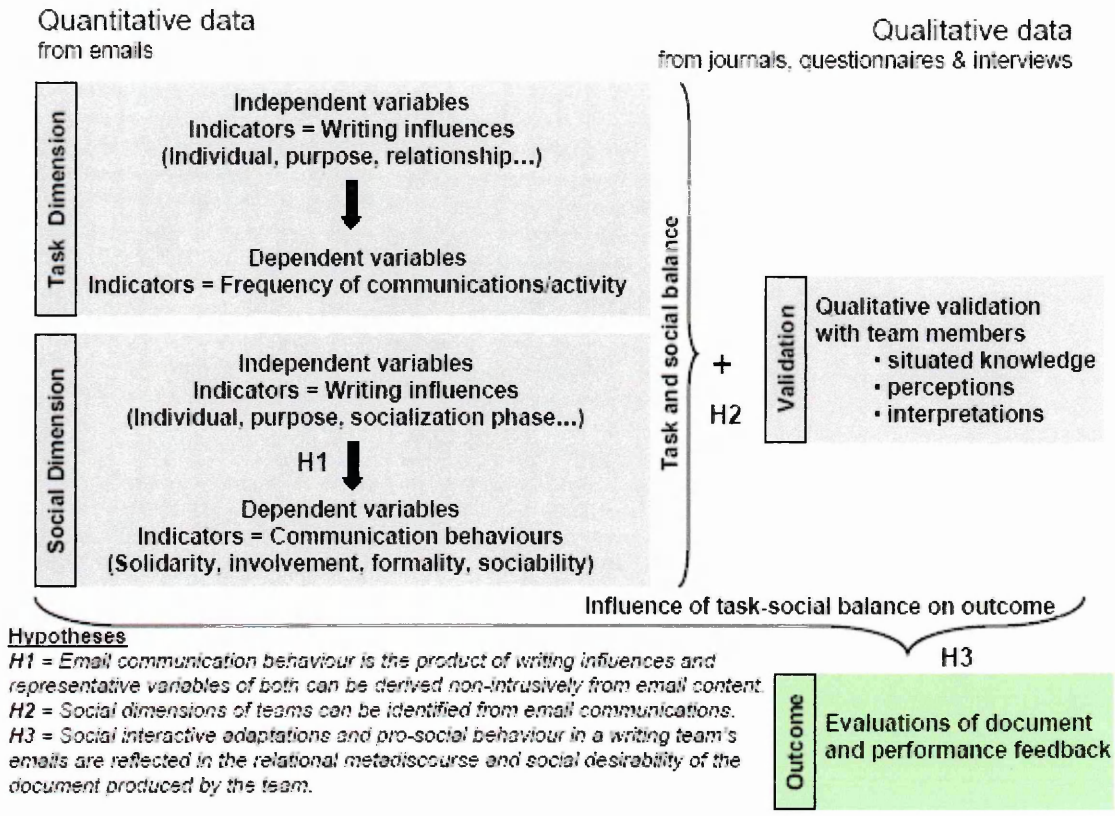

Figure 7-1: Research Framework highlighting H3 phase

This research aims to answer the research question "Can we learn about the influence of team culture on virtual team writing from content analysis of email communications during projects?" I describe team culture by profiling the balance between social and task dimensions on projects. The previous studies reported in chapters 4, 5 and 6 showed that indicators of influences on writing could be derived from email records and shown to influence social behaviour in email communications in networked team writing projects. Social interactive adaptations in communication behaviour with writing influences helped to describe the balance of the social and task dimensions on the projects. There is strong evidence in the literature that a social dimension in team work contributes in a positive way to the performance of the team (see under "Social dimensions and team performance" in section 2.3.2). The balance between the social and task dimensions on a project therefore affects the team performance. The aspect of performance under focus in this chapter is the social desirability of the final documents. I evaluate documents and available feedback on the documents from commercial and academic

team writing projects. I then review these findings together with the balance between task and social dimensions of these projects to address the following hypothesis:

> *H3 = Social interactive adaptations and pro-social behaviour in a writing team's emails are reflected in the relational metadiscourse and social desirability of the document produced by the team.*

This part of the overall research project is shaded green in Figure 7-1, which shows how this part of the research fits within the overall framework. The Keyware project reported in chapter 4 was a pilot study and is not discussed in this chapter.

The commercial (Namahn) project produced Owner and Service manuals for the installation of a payment terminal in shops, and the academic (Society) project produced the handouts for a postgraduate clinical training course. There were six main team members on the Namahn project and eighteen main actors on the Society project. Namahn produced the documentation for a client, Banksys, in a commercial context, and the Society produced the handouts for course attendees in an academic context. The Namahn project had a finite length, running for two months, albeit in a long-term supplier-client relationship. The Society project has been ongoing since 2002. Further details on these two projects are presented in sections 3.4.3 and 3.4.4.

In section 7.2, I evaluate the documents which were produced in these projects. To evaluate the documents in a systematic way for comparison relevant to the social dimensions of working, I use Sless's (2004) social desirability model. Expert evaluation against Sless's social desirability criteria provided a method for comparing the documents, taking into account the different audiences, purposes and contexts of use for the documents.

De Jong and Schellens (1997) argue that "although text-focused and expert-focused methods may provide valuable feedback on documents...they cannot replace reader-focused evaluations" (de Jong and Schellens 1997 p403). Research by Lentz and de Jong has demonstrated the limitations of expert-focused evaluations. In reporting their research, they acknowledge that generalising results from the study of a small number of texts using one reader-focused and one expert-focused method requires caution. However, they found that technical writers predicted less than 15% of the problems identified by readers, and also detected many new problems, which they labelled as "false alarms". The researchers point out, however, that "such wrong predictions can still be useful for the revision" (Lentz and de Jong 1997 p228). Additionally there was little agreement between technical writers on problem detections: "experts tend to mix personal biases with assumptions about readers' behaviour" (Lentz and de Jong 1997 p232). These limitations need to be taken into account when considering the results of an expert-focused document evaluation as used in this research.

Reader-focused methods, such as user protocols or focus groups, were not possible within the scope of this research. Permission was kindly given to interview end-users (shop keepers) of the commercial documents in Belgium, to gather information on real use of the document. However, the new version of the document was not yet in use and when eventually used, would

incorporate amendments made by distributors, thus not directly reflecting the team culture studied in this research. Due to the limitations of expert evaluation of documents (Schriver 1997 p473), I also discuss user and interviewee feedback in section 7.2.4.

To compare the relative task and social components, I use social marker frequencies, correlations between social markers, and task profiles of the two projects. I also review the conclusions from the case study analyses reported in chapters 5 and 6. In these studies I evaluated the social-task balance in projects based on adaptations of communication behaviour and indicators of group cohesion. These social-task balance comparisons are presented in section 7.3.

Finally, in section 7.4, I use the combined data collected to assess the relationship between the social-task balance on projects and the project outcome in terms of the document evaluations and feedback from the project leaders for the documents. This analysis thus explores the concept that social interactive writing behaviour demonstrated in socio-emotional behaviour in team emails is reflected in the final document.

7.2 Document evaluations

7.2.1 Overview

The concept I explore in this part of my research is whether the way teams adapt to meet social and task needs in communications exchanged during the writing project (which also describes the team culture) is reflected in the relational metadiscourse and social desirability of the final document. To evaluate the two documents produced in the commercial and academic writing projects using a standard framework, I used Sless's social desirability model. The focus of this model on the social dimension parallels the focus of my research on the social dimension of working, to explore the influence of the social-task balance or team culture on virtual team writing (see section 2.3.2).

The social desirability model was designed by Professor David Sless, who defines information design as "concerned with making information accessible and usable to people" (Sless 1990 cited in Carliner 2000 p564). Sless set up the Communication Research Institute of Australia (CRIA) in 1985 and is currently Vice President of the International Institute for Information Design. With over 180 publications, he is a member of the Editorial Board of the Information Design Journal and is currently researching regulations and guidelines for effective communication (CRIA 2005; David Sless n.d.). He and his colleagues at the CRIA have been instrumental in setting socially acceptable standards for document design, which in 2003 they specified as being able to find and use 90% of what is being searched for (Sless 2004 p35). Methods used by the CRIA over the last 20 years have been applied to government, IT, healthcare and many other documents used in the public domain for communication between organizations and individuals (CRIA 2005; David Sless n.d.). Sless's approach captures much that is leading edge in the field of information design and is therefore used worldwide.

At Sheffield Hallam University we have been using Sless's model as a teaching tool to help M.A. students, who are professional technical writers, evaluate and understand the social interactive role of written documents. Student evaluations of this exercise over the last two years have consistently demonstrated the value of the method. The model also fits with my own understanding both of written communication as a meeting of writer and reader minds, and of the distinction between spoken and written communication: "written texts are composed for a context of eventual or potential use... determined by time,.... place...and purpose" (Nystrand 1986 p95). Document evaluation based on Sless's definitions of social desirability provides a social interpretation of quality, fitting with my own research stance and more directly relevant to the social dimensions of working.

Sless emphasizes that writing and designing are not separate tasks but "go hand in hand" (Sless 2004 p33):

> *We need to bear in mind that behind this achievement [improved documentation] lies the exercise of traditional document design skills, without which this level of performance would be impossible. To achieve these results requires the very best of imaginative typographic design and writing, plus the rigorous disciplines of editing and graphic refinement that are a necessary part of traditional good document design.....It is the way in which the old [traditional skills] and the new [social desirability] are unified that makes these results achievable in practice (Sless 2004 p34).*

This bridge between traditional document design and social desirability justifies the value of text-based evaluation of documents by editorial review. Preston describes heuristic review as "using a predefined set of standards, a professional usability expert reviews someone else's product or product design and presents a marked checklist back to the designer" (Preston 2004 p15). Many frameworks and checklists of design elements are available in the literature for document evaluation and editorial review (e.g. Barker 1998 p243; Carliner 2000 p564-570; Porter and Coggin 1995 p261; Rude 2002 p240; Van Buren and Buehler 1980 p13-23). Sless first approaches document design or improvement by researching what people want to do with documents. "By focusing their attention on the outcomes rather than the content, we are deliberately distancing them from the writing, design and editing of the document" (Sless 2004 p28). Thus while recognizing traditional information design skills, Sless is also advocating a starting point from a slightly different perspective, that of social appropriateness. Rather than using a checklist of document design attributes, therefore, I address the questions which Sless asks related to each of the seven attributes of his model and explain my responses in terms of the relevant design attributes. These evaluations, therefore, are not a systematic review of all document design attributes, but focus only on those which highlight issues relevant to social desirability of the documents under review. Where a particular design attribute (e.g. typography or tone) is relevant to the social desirability of one of the documents, I have also commented on that aspect for the other document.

Sless's model addresses seven different aspects of a document: credibility, respectfulness, attractiveness, physical appropriateness, and whether the document is efficient, easy to use and productive. These attributes are often difficult to distinguish between in evaluating documents.

Carliner points out the overlap among levels in his own model of a three-part framework for information design (physical, cognitive and affective); "... clear distinctions among the different but related issues do not always exist" (Carliner 2000 p570). Indeed this difficulty of distinguishing between information design attributes is also reflected in definitions of levels of edit. In reviewing Van Buren and Buehler's levels of edit, Haugen discusses the difficulty distinguishing between substantive and surface-feature edits (Haugen n.d. p7). In spite of this overlapping between the levels of different frameworks used to design or evaluate quality of documents, the value of using a framework lies in drawing attention to relevant issues in a systematic way.

My evaluations are limited to the perspective of an editor, so I cannot comment on the accuracy of the subject matter content. (For my own experience and qualifications relevant to editorial review please see appendix H). However, data gathered from the six main team members on the commercial project and the interviewee on the academic project provide some indication of the level of subject matter expertise (SME) available. On the commercial project, three of the team members had relevant subject matter degree qualifications; one had SME experience of more than 5 years, one between 2 and 5, and three had less than 2 years. On a scale "very low", "quite low", "don't know", "quite high" and "very high", two participants rated their SME knowledge as "quite high", two "quite low" and one as "don't know". On the academic project, the Course Leader rated team members on a three-point scale of "high" "average" or "inexperienced" professional expertise (appendix K line 45-46). She rated 16 of the 18 main team members as having "high" professional expertise, and the remaining two as having "average" expertise. Additionally she commented:

> *We wouldn't have invited them to participate if they weren't experts in their fields (appendix K line 39).*

There thus appears to have been strong subject matter expertise available on both projects and both had peer level review processes to control for subject matter accuracy.

In the following section I describe what people want to do with the documents and then in section 7.2.3, I compare their social desirability. In section 7.2.4, I include feedback on the documents available from the client in the commercial project, and from end users in the academic project.

7.2.2 What do people want to do with the documents?

As a measure of team writing performance for the commercial project, I am evaluating the C-ZAM/XENTA Owner and Service Manuals, mainly concentrating on the Owner's manual. The primary audience for the documents is international distributors, who then adapt versions for local use (and language). Thus the primary audience acts as an intermediary adapting the manual for end users. The end users are either distributors' engineers or shop keepers, who install the C-ZAM/XENTA unit in shops. In either case, therefore, the manuals are used in shops to install a piece of hardware. The Service manual is used by an engineer, i.e. someone with

more knowledge of hardware, and the Owner's manual is more likely to be used by a non-expert. Although the two manuals require more or less the same content, the shop keeper's (Owner's) version required a slightly different tone. For the Owner's manual, the relationship between the organization issuing the document (Banksys) and the end user is a supplier-client relationship. The end user needs clear safe instructions on how to install the unit in a shop and what do in case of problems. Based on her own research and a survey by DeTienne and Smart (1995 cited in Schriver 1997 p214) showing that users do read hardware manuals, Schriver writes "Given the fact that consumers do read manuals, companies might consider the manual as a good way to help reinforce a positive corporate identity after the sale of a product or service has been made" (Schriver 1997 p214). The supplier thus needs to provide the information required to complete the sale, satisfy the customer and maintain credibility of image to retain customer loyalty.

The content and design of the two manuals are almost identical, with some additional tasks explained for engineers in the Service Manual. The Owner's manual is about 40 pages long and the Service manual about 50 pages. These documents have up to eight main sections, divided into subsections of no more than two or three pages.

As a measure of the team writing performance for the academic project, I am evaluating the course handouts, or lecture notes. The course content is represented in these documents, which the Society aims to publish in a handbook. The Society is the organization providing the subject matter expertise via the handouts, so that the relationship between the issuer of the document and the reader is a teacher-student relationship. Course attendees are all practising professionals or researchers in the clinical topic, but vary in their particular expertise, including for example surgeons, technical engineers and others. The documents need to provide forms for processing workshop data and supporting information for the FtF teaching over a three day period, so that students can concentrate on the presentations rather than writing copious notes. Students need to annotate handouts with personal notes and reminders, and should be able to use the materials later for reference. Lecturers need handouts to refer to during teaching, either to refresh themselves on what they plan to cover, or to explain and support particular points during their teaching.

The Society produced three sets of handouts, one for each of the three days of the course. The handouts for each day are 80, 65 and 65 pages in length. There were three workshops per day, and between four and six lectures per day. The handout for each lecture is approximately 4 to 6 pages long and for the workshops up to 30 pages each, including tables for recording results.

7.2.3 Comparison of social desirability evaluations

Overview

As described above in section 7.2.1, I considered each of Sless's social desirability attributes for each of the document types. Below, I explain the attributes and justify my evaluations. Where a

particular design attribute (e.g. typography or tone) is relevant to the social desirability of one of the documents, I have also commented on that aspect for the other document. Table 7-1 on page 214 summarizes the issues which emerged from this analysis.

Credibility

To be credible, a document has to present an authoritative stance: "it must not only be accurate and authorative but must be perceived as accurate and authorative" (Sless 2004 p26).

In the manuals, the logo and product name on every page adds credibility, presenting the corporate identity of the producer and therefore the authority on the product to be installed. Copyright and confidentiality notices on the first page confirm ownership of intellectual property related to the product to be installed, further reinforcing the issuing organization's authority. Additionally, the appropriate relevant authorative sources are referenced for safety recommendations and regulations. While the layout is consistent, presenting a professional image, there is some scope to improve small grammatical errors and awkward phrasing, to improve readers' perception of the documents' accuracy.

Society representation is very low in the handouts. There are no logos or copyright notices on the handouts or folder. Affiliations and credentials of author-lecturers, which would have increased the academic credibility of the documents, are not included. Small inconsistencies in heading levels and use of typography for key words etc. also suggest a collation of contributions from different research and clinical organizations, rather than a coherent delivery from the Society. "Companies convey a sense of stability when all documents look as if they belong together" (Hackos 1988 cited in Nord and Tanner 1993 p224-225). Correct grammar and natural phrasing, however, lend credence to the documents' accuracy.

Respectful

Sless gives examples to demonstrate whether documents show readers respect (Sless 2004 p26). These examples highlight attributes from which readers develop a perception of whether the document issuers really care about them and what they need from a document.

Concise, well-focused content with little redundancy in the manuals supports clarity and readability (Kirkman 1992 p17) and allows readers to become productive quickly (Carroll 1990 cited in Carliner 2000 p563). This shows respect for busy shop keepers, whose objective is to install the unit safely, correctly and as quickly as possible.

In the academic documents, concise handouts also without redundancy support clarity and readability, to help understanding and learning during and after the course. Course attendees would be left with no doubt that the course content had been well prepared and that their need to concentrate on oral presentations rather than taking notes had been taken into account.

Language tone could be fine-tuned in the manuals, to emphasize the organization's respect for the customer. The first page of the manuals protects the liability of the supplier with an entire paragraph in capital letters and a potentially offensive tone, using phrasing such as "SHALL

NOT BE LIABLE" and "is strictly forbidden", and legal-type terminology such as "contained herein" and "construed or interpreted as an implied obligation". These notices could be delivered in a simpler and more respectful way. "Plain English can replace legal language without any loss of certainty or accuracy" (Plain English Campaign 1993 p46). Other warnings are expressed elsewhere in the document to protect the supplier and could also be reworded in a more respectful way for the customer. Additionally in the manuals, a page is dedicated to instructing on how to tear off tickets, with four illustrations. This content could be patronising to shop keepers or engineers accustomed to tearing paper receipts off similar devices.

The tone of language used in the handouts is not patronizing. Short sentences use the necessary clinical terminology, but present information in plain English. The design of the course builds from basic theory and terminology towards application, clinical assessment and more complex topics. In this way the course content does not assume expert knowledge from the beginning. Students came from different professions within the same clinical field (e.g. surgeons, technical engineers, etc.). Although this may have caused some mismatch with expert needs and potential for patronisation at the beginning of the course, it catered for others who needed priming on the basics, thus appropriately meeting the needs of the broad audience.

Attractive

> *Readers' willingness to use a document (printed or online) [is] based on its appearance...balanced page layout, interesting graphics, and legible print invite readers to use documents...Attractive documents look as if the producers value the information enough to care about its appearance and suggest that the same care has gone into writing the text as was invested in its design (Nord and Tanner 1993 p224).*

Readers interpret high quality presentation as representative of high quality text and are thus motivated to read documents.

> *One way of bestowing importance, value and dignity to the reader's task is to make documents attractive...It is about long-term satisfaction, so that each communicative occasion enhances the relationship between the reader and the document (Sless 2004 p26).*

Attractiveness thus encourages readers and helps to sustain long-term satisfaction.

Both documents have a relatively large and clear typeface (>2mm x-height) to help legibility and attractiveness (Plain English Campaign 1993 p56; Schriver 1997 p258-259). However, the table of contents and some of the liability notices in the manuals are in capital letters, which is less attractive and also more difficult to read (Breland and Breland 1944 cited in Schriver 1997 p274; Rude 2002 p148).

The manuals have text aligned with ragged right edge. The academic document has justification of text against both left and right margins, causing irregular word spacing across the page. This detracts from the appearance and causes uneven spacing between words which makes text harder to read (Burnett 1997 p211; Plain English Campaign 1993 p62; Schriver 1997 p270).

Pictures are more quickly understood, are remembered better than text, and also attract readers (Horton 1993 p191). Illustrations supplement the text to improve understanding and retention (Levie and Lentz 1982 cited in Schriver 1997 p408) in both the manuals and the handouts. They are uncluttered (Hartley 1994 p108; Tufte 1997 p48) and located appropriately with the text they support (Hartley 1994 p82, 108), thus optimising comprehension. In the manuals, they help to convey information, support the text, and instruct readers on how to complete the tasks (Rude 2002 p331). In the handouts, illustrations help to convey information and support comprehension of the text (Rude 2002 p329).

Use of colour is "an especially appealing device often contributing significantly to the effectiveness and clarity of a document" (Burnett 1997 p221); both documents use some colour in illustrations. Both documents also use plenty of white space in margins, leading and to frame illustrations, which helps reading and assimilation of information (Barker 1998 p308; Smith and McCombs, 1971 cited in Schriver p275) and also makes documents more attractive (Burnett 1997 p210; Strong 1926 cited in Schriver 1997 p275).

Typography distinguishes section headings providing visual cues to the organization of information (Burnett 1997 p610; Rude 2002 p 315; Schriver 1997 p284) in both documents and additionally to distinguish lists, and keywords in the academic documents. There is, however, some inconsistency in use of typography in the handouts, which may confuse readers and detract from learning and retrieval. "Consistency in type style and placement of headings helps readers perceive the organizational pattern. Inconsistencies may suggest changes in meaning where no change was intended" (Rude 2002 p126).

Physically appropriate

A physically appropriate communication is delivered in a medium appropriate for the context of use (Sless 2004 p26). The manuals are delivered appropriately on paper, to use in the shop while installing the terminal and for future reference with troubleshooting. Similarly the course handouts are provided on paper in a folder, in the sequential order of presentation of the lectures, for use during and reference after the course.

Socially appropriate

> *Building a positive identity (and here I am talking about more than just logos, product naming, or graphic style) requires organizations to develop a distinctive voice —through the interplay of text and graphics— that makes evident to audiences that their knowledge and values are understood, respected, and not taken for granted (Schriver 1997 p204).*

Paralleling the approach of studying how writers adapt their emails to their readers, Sless writes that readers "judge organizations on the extent to which those organizations match or adapt to their expectations" (Sless 2004 p27).

The social relationship between the message sender and receiver for the Owner's manual is a supplier-customer relationship. At the beginning of the documentation project, a client team member had recommended including promotional material in the manuals. The supplier

recommended against this (appendix L, point 7). On the one hand, shop keepers are busy and want clear concise instructions to accomplish the installation task, rather than paragraphs of promotional narrative telling them what a good product they have bought. On the other hand, some way of showing appreciation for customer loyalty is appropriate in this particular relationship. The template could have included an optional sentence for local distributors to adapt as appropriate, thanking the customer for buying the device and providing a contact number for customer service. Showing appreciation of the customer without including promotional content has the potential to improve how readers relate to the document, thus building on the corporate identity of the document issuer and helping to maintain the supplier-client relationship.

The relationship between the Society and the course attendees is a teacher-student relationship, with experts sharing knowledge with other professionals working in the same clinical field. Thanking course attendees in the lecture handouts would be contextually inappropriate, as these documents are delivered in a FtF scenario with a welcome address, in which the Chairman thanks attendees personally.

The tendency to write in a negative way, which is more difficult to understand (Plain English Campaign 1993 p32; Rude 2002 p262), and with longer or less familiar words (Kirkman 1992 p18-24; Turk and Kirkman 1989 p100-104) contributed to an impersonal tone in the manuals. An example of negative phrasing is *"make sure that the X is not powered on".* Using a positive valence, such as *"check the X is switched off"* improves accuracy and speed of understanding (Rude 2002 p262), and is also more direct and appealing (Burnett 1997 p294). Lengthy phrasing such as *"the X is an optional accessory designed for the merchant who does not use a Y",* could be simplified with a more conversational style, such as, *"the X is an alternative to the Y". Other examples are:*

> *...breakdowns or any anomaly that may occur due to incorrect manipulation or use of the terminal.*
>
> *verify that the surface upon which you are going to attach the fixation plate is flat*
>
> *The fixation plate enhances the stability of the...*

In contrast, apart from the necessary terminology, author-lecturers used plain English in short sentences with a simple, more personal and therefore more effective communication style in the handouts (Barker 1998 p331-332; Turk and Kirkman 1989 p112), e.g.

> *Now that we understand A and B, how do we define C?*
>
> *Understand your measurement system and its limitations.*
>
> *Check for accuracy and alignment with the camera system.*
>
> *Take measurements in the context of the clinical problem.*
>
> *Here we are in a situation of equilibrium.*

Efficient

To define the efficiency of a document, Sless asks "how long does a person have to spend looking through a guide to a particular service to find what they are interested in?" (Sless 2004 p27).

Informative section headings improve reader focus, readability, information retention and retrieval (Burnett 1997 p212; Nord and Tanner 1993 p222- 225; Rude 2002 p315) both in the manuals and the handouts. However, navigation in both documents has some scope for improvement.

Cross-referencing is a device to help readers locate information (Burnett 1997 p427) and is missing in both the manuals and the handouts. In the manuals, for example, suggestions are provided in the troubleshooting section to help the shop keeper with problems. Some of these solutions ask the reader to complete an action which has already been explained elsewhere in the text, such as "Check the A cable is connected to the B outlet." However, without explicit cross referencing to the full instructions and illustrations of "A cable" and "B outlet", shop keepers will need to refer back to the table of contents, or to rifle through the entire document to find the appropriate instructions. Cross referencing, to guide the reader directly back to where these instructions are in the document would save the reader time and frustration. The section headings, therefore, need to be numbered and this would also help distinguish between them and between different levels of sections. An index of key terms would also help information retrieval (Burnett 1997 p427; Nord and Tanner 1993 p222), for example, if a reader wanted to search for the instructions about a particular item, such as a 'security chip'.

Similarly, the course handouts from the academic project require section numbering, which would allow cross referencing, and an index which would help attendees to use the information for reference after the course. Lack of section numbers, a table of contents or index would impede navigation when using the materials for reference after the course, and may even have hindered navigation during the course (one respondent did comment on this; see Table 7-2 on page 216).

Ease of use / usability

Sless describes "ease of use" as a usability measure, asking "Can they [users] find what they are looking for? Can they understand it when they find it? Can they use the information appropriately? Is there enough information for them to use it appropriately?" (Sless 2004 p27).

In both the manuals and the handouts, content is organized into sections, making information easier to understand and remember (Rude 2002 p287).

There is some scope to improve ordering of content in the manuals. Safety recommendations relating to different activities such as installing and opening units are located at the end and could be more usefully located both at the beginning and in the appropriate sections on installing and opening the units, to avoid the reader having to move backwards and forwards in

the document (Burnet 1997 p431). Some warnings following instructions within sections need relocating to before the instructions to help avoid errors. Ordering of content in the handouts is logical, building progressively from simpler to more complex topics, which supports learning (Rude 2002 p291).

Within sections of the manuals, there is also some ambiguous content and some illustrations have no labels, making the instructions difficult to understand. For example instructions on installing a printer (in the Service manual) refer to "printer connectors", "cable locks", and "printed circuit boards", without any of these items being labelled in the associated diagram. Where illustrations are labelled, items are labelled by numbers which cross-reference to tables listing the item names. Direct labelling would be quicker to read and assimilate, and also take up less space. In the handouts, content is unambiguous and illustrations are labelled directly.

There is potential to improve usability of both documents with a little redundancy. In the manuals, for example, the product and its main accessories are shown and their parts itemized at the beginning. Other items are introduced later so that it is difficult to conceptualize how they all fit together. A complete overview of how all the parts fit together, what they do, and choices which need to be made up front might help the reader to create a mental overview of the tasks to be done, before moving on to the individual steps required for each part of the process. This would provide what Burnett describes as the "part/whole organization" overview (Burnett 1997 p180) and what Rude refers to as the "macrostructure" (Rude 2002 p287, 291). This overview section could also cover what decisions need to be made before installation, and what additional equipment might be needed. For example, I would prefer to know before I start that I'm going to need a "calibrated torque screwdriver with setting 0.7Nm and a T 10 Torx head". Knowing before I start that I can choose to run cables vertically downwards under the unit through the counter surface, or simply horizontally through the back of the unit allows me to prepare my counter surface as necessary. These small details on the screw driver needed and choice of cable positioning are introduced with the detailed set of instructions for installation of different parts, and might be better covered in an introductory section titled "What you need to know before you start".

Overview and summary information for each day's handouts would also improve coherence across the handouts, helping to build a mental overview of the whole course. Additionally, overviews and summaries within each handout would help to build a mental overview to support learning for each lecture (Rude 2002 p287, 291).

Some of the words and phrases in the manuals, such as "depicted", "optimal functioning", "verify the functionality" and "facilitate the merchant's tasks" could be shortened and replaced with more familiar words, both to help readability and local translators (Kirkman 1992 p153; Turk and Kirkman 1989 p99). Some abbreviations are not explained and some of the less common terms, such as "bevel" and "strain relief" could be replaced with words more likely to be intuitively understood by shop keepers. If these words do need to be used, a glossary could be included

with explanations of what they mean. In the handouts, terminology is explained when first used, although some abbreviations are not explained.

Further in the manuals, as an additional courtesy to the shop keeper to make further actions easier, the actual contact details could be included wherever a recommendation is included for the reader to contact the supplier. Additionally, concrete rather than abstract external references to other documents would be more informative. Explicit referencing to external documents is included in the handouts, although there are no explicit contact details for the Society or author-lecturers.

Productive

Sless also argues that "people need to find their engagement with documents productive...the information needs to be useful, reassuring, or leading to something new" (Sless 2004 p27).

In the manuals, the ergonomics recommendations for where to install the unit provide a particularly good example of content supporting productivity. These recommendations show an excellent appreciation of the actual context of use of the device and needs of end users. For example, they take into account wheelchair shoppers and all the different things shop keepers might need to do with the terminal, such as changing the paper in the printer, and the space they will need for this. This section relates directly to the reader and the reader's context in a very practical way. The manuals lead to the installation of the C-ZAM/XENTA, thus supporting the implementation of a new tool in the shop keeper's working environment.

Course attendees attend the course to increase their knowledge for their professional practice or research. The course handouts provide a tool to reinforce their learning both at the time of the course and as reference material for after the course. References for sources cited during the lectures and workshops are also included in the handouts, so that attendees can study further after the course.

Review of document evaluations

Table 7-1 summarizes evaluations of the social desirability attributes of the commercial and academic documents. Comparing evaluations, both documents are highly productive, physically appropriate and attractive, although there is some scope for improvements to support ease of use and efficiency in both, for example by improving navigation with section numbers etc. The strongest attributes of the course handouts are their social appropriateness and respect for the reader; the weakest attributes relate to credibility and efficiency. The manuals have strong corporate credibility, with scope for improvement in social appropriateness and respect for the reader. The issuing organization has a higher profile in the commercial than in the academic documents. Together with the high strengths of the course handouts in terms of social appropriateness and respect for the reader, I conclude that the course handouts show a more even balance than the commercial manuals, between the task and relational elements communicated to readers.

This conclusion requires consideration of three limitations to the approach used. The first is the influence of my own personal perspective on the results of the evaluation, as the researcher. In making judgements on the documents within the framework of Sless's social desirability model, my own biases may have confused assumptions I have made on the readers' behalf: "experts tend to mix personal biases with assumptions about readers' behaviour" (Lentz and de Jong 1997 p232). The second limitation is the variation in relative severity i.e. the "damage that the problem might cause for the document's effectiveness", (Lentz and de Jong 1997 p 227) of the items listed in Table 7-1. No estimations of the severity have been made in this research. Readers may be less concerned about a missing logo than about patronising content. We may not assume, therefore, that all the items listed are equally important to the document's social desirability. Finally, as discussed at the beginning of this chapter, expert-focused evaluations have several limitations and cannot replace reader-focused evaluations, so that conclusions drawn on the basis of the results of my evaluations require caution.

Table 7-1: Social desirability comparisons of commercial and academic documents

✓= strength; ! = potential for improvement; × = weakness; ? = unknown; ▪ = not applicable

Attribute	Commercial	Academic
Credibility	✓ logo and product name on every page ✓ other relevant authorative sources ✓ consistent layout and typography × some grammatical errors and awkward phrasing	× no Society logo, no affiliations or credentials of lecturer-authors ✓ other scientific sources cited × inconsistent typography ✓ correct grammar and natural phrasing
Respectful	✓ concise information with no redundancy × legalese tone, some potentially offensive typography × some potentially patronising content	✓ concise information with no redundancy ✓ Respectful tone ✓ All content appropriate for broad audience
Attractive	✓ black and white with some use of colour ✓ clear illustrations ✓ easy to read font ✓ plenty of white space ✓ typography used to distinguish headings, lists, and keywords etc ✓ consistent typography × inappropriate use of capital letters	✓ black and white with some use of colour ✓ clear illustrations ✓ easy to read font ✓ plenty of white space ✓ typography used to distinguish headings, lists, and keywords etc × inconsistent typography × inappropriate justification
Physically appropriate	✓ paper manual appropriate for context of use	✓ paper handouts in a folder appropriate for context of use
Socially appropriate	× no acknowledgement of appreciation for customer loyalty × impersonal tone, formal style	▪ inappropriate in handouts ✓ personal tone, conversational style
Efficient / navigation	✓ table of contents and page numbers × missing section numbers and cross-referencing × no index	× no table of contents or page numbers × missing section numbers and cross-referencing × no index
Ease of use / usability	✓ organization of information in sections ! some content inappropriately ordered ! some ambiguous content ✓ clear illustrations × indirect labelling of illustrations ! some missing illustration labels × missing overview ! some difficult to translate phrases and unexplained abbreviations ! some unexplained terminology with no glossary × abstract referencing to documents × abstract referencing to contacts	✓ organization of information in sections ✓ content appropriately ordered ✓ understandable content ✓ clear illustrations ✓ direct labelling of illustrations ? [unknown whether missing labels] × missing overviews ! some unexplained abbreviations ✓ clinical terms defined when first used ✓ explicit referencing to documents × abstract referencing to contacts
Productive	✓ supports the implementation of a new tool ✓ provides troubleshooting reference material	✓ increases professional knowledge ✓ provides study reference material

7.2.4 Feedback on the documents

Overview

Comments provided by the client interviewee and from the client debriefing on the commercial project are presented in the next section. This feedback is provided as an alternative to end user feedback on the manuals, which would not reflect the quality of the commercial project, due to rewriting by distributors. User feedback provided by the Course Leader from the 2003 course is presented for the academic project in the subsequent section.

Commercial document: interviewee feedback

Both the supplier and client interviewees in the commercial case study commented on the quality of the finished product. When asked how well the document had met their needs, the client interviewee responded:

> *For us there were no problems (appendix O, line 116).*

Further I queried what kind of feedback they had received from distributors using the documents, to which the interviewee responded:

> *We don't have any real feedback. In Belgium they have used their original way of documenting, but we have now asked them to stick to what we have offered. We don't have any feedback from customers either. If something needed to be changed we would probably hear this, but most changes come from us (appendix O, line 119).*

From the debriefing which Namahn presented to Banksys, it is clear that the manual was what the client had expected, providing a starting point for their distributors' documentation, covering appropriate liability issues, and presenting high quality illustrations. The Namahn interviewee reported, however, that Banksys were slightly disappointed with the terminology used in the manual (appendix L, point 3). There was thus some scope for improvement in the commercial documentation, although the client and supplier satisfaction appeared to be high.

Academic document: user feedback

Evaluation forms completed by end users of the Society case study were collected by the Course Leader from 12 (of the 50) course attendees after the 2003 course. Feedback relevant to the handouts, as opposed to FtF delivery of the course, are summarized in Table 7-2. User evaluations of the handouts, content relevance and applicability to professional practice were all high. Appropriateness of the level of the course was rated lower. One respondent commented that this was unavoidable due to the heterogeneous nature of the student group in terms of background, education and experience (see Table 7-2). In the post-analysis interview, when the Society interviewee was asked to compare the performance of the project to other similar projects, she commented:

> *For comparison with other projects – there aren't that many other projects that deliver the same. Last year we had a grading of 4 to 5 on everything on the course, so that's good (appendix K, line 105).*

The uniqueness of the postgraduate clinical training course means that the students who attend have different levels of expertise and different needs, which was reflected in the low score for appropriateness of the level of the course material in the 2003 feedback. The scores of 4 to 5 in the 2005 feedback suggest that this aspect was improved on for later courses. Further in the evaluation questionnaire in 2003, respondents were asked whether the course met their expectations. 11 out of the 12 attendees replied positively; one did not respond to this question.

Table 7-2: Society course user feedback

Aspect	Mean score and course attendee (n=12) feedback (1=very poor 5 = very good)	
Handouts	4	▪ Could improve with more information in more detail, must be numbered and titled on every page. ▪ Too much paper. ▪ A bible, with some small deficits. ▪ "A lot of paper came later, or was additional [addendums]" ▪ "....remember the language! Not all of us know right away what, for example, the abbreviations mean, ..." ▪ "Just a little more time needed checking the handout material for faults, twisted numbers R/L, missing minus signs etc." ▪ "You could improve on the quality control for the handouts and graphs." ▪ "The interpretation of data is in itself difficult. For this reason info given from clinical examination MUST be impeccably right."
Content relevance	4.3	
Applicability to practice	4.0	▪ "More in depth details towards applicability necessary" ▪ more on ..."
Appropriateness of level of course	2.8	▪ "Unavoidable when the group is heterogeneous in background, education and experience."
Missing content and suggestions?		▪ 5 suggestions for additional content ▪ 1 suggestion for an advanced course ▪ 3 positive comments
Quality of the Faculty		▪ 7 positive comments
Did the course meet your expectations?		▪ 11 positive comments

Relevant to the handouts, one attendee commented on the difficulty understanding abbreviations when English was not a first language. Two attendees referred to proof reading quality control; two commented on the amount of paper (without the addenda the handouts for the three days totalled 210 pages), and one pointed out the difficulty navigating the documents without page numbers and titles. One attendee also commented on the necessity for accurate technical information, although it is not clear whether this comment was a general observation, or related to a particular error in the course material. There were several suggestions for additional content, although these were invited, not given as justifications for poor ratings.

The high ratings for the handouts, content relevance and applicability to practice, and in particular the positive comments for the last three questions in Table 7-2, show that the overall impression of the course and the handouts was high. Although there was scope for additional course content, the end users were satisfied with the current quality. There was scope, however, to improve navigation, for minor errors to be corrected, and full explanations of abbreviations of English words would have been helpful for non native English speakers.

Review of feedback on documents

Satisfaction in the documents was thus high on both projects. End users were satisfied with the quality of the handouts, although there was scope to improve navigation and correct minor errors. Clients in the client-supplier relationship were satisfied that the manuals met their specifications for a starting point for distributors' documentation, including liability disclaimers and high quality illustrations. Although the client was disappointed in the terminology, overall satisfaction was high.

7.3 Social and task dimensions

7.3.1 Social marker frequencies

Comparing social markers between the two projects shows significant differences between elaboration, open greeting length, formality, involvement, solidarity and sociability at $p < 0.05$. The academic context had higher means for all these markers (see Table 7-3). Close greeting length did not vary, suggesting that close greetings are subject to more pervasive social norms (thus operating across projects) than open greetings. Interestingly, however, close greeting behaviour correlated with all the social behaviour markers (involvement, solidarity and sociability) in both projects.

In particular, the differences in mean marker frequencies for involvement, solidarity and sociability suggest that these feelings were represented more in communications on the academic project than in the commercial project. I therefore conclude that the academic project communications showed higher socio-emotional content.

Table 7-3: Results for Wilcoxon-Mann-Whitney Two Sample test (for all emails on both projects with more than 5 words in the body text)

	Body text word count	Open greeting word count	Close greeting word count	Standard-ised formality score	% first person singular pronouns	% first person plural pronouns	Social building units/ email
Pr	<.0001	<.0001	<.1875	<.0001	<.0001	<.0326	<.0001
Commercial project means	62	1.3	1.5	5.0	3.4	1.0	0.6
Academic project means	111	1.9	1.7	5.4	4.5	1.1	1.0

7.3.2 Correlations

In section 3.9 on "Correlations between social dimension markers", I introduced the idea that differences in significant correlation coefficients between the communication markers might indicate differences in behaviour on the two projects, and I return to this point here.

Firstly, involvement and sociability had a small positive correlation for the Namahn project ($r = 0.18$) and larger correlation for the Society project ($r = 0.49$). The involvement marker represents involvement in the task and the socio-emotional content of the message, so this difference could reflect that although the task component is represented in the Namahn data, a lower socio-emotional content is represented. On the Society project, sociability also correlated significantly with solidarity ($r = 0.24$), and solidarity with involvement ($r = 0.35$). These relationships were not significant for the Namahn data; thus in Namahn communications, feelings of solidarity did not vary with feelings of involvement or use of social building strategies, whereas this was the case in the Society project.

Longer emails showed higher involvement ($r = 0.36/0.72$ for Namahn/Society) and solidarity ($r = 0.53/0.70$ for Namahn/Society) in both projects. However, there were no significant correlations between body text length and sociability in the Namahn project, whereas in the Society project, longer emails also showed higher sociability ($r = 0.43$). This difference between the two projects suggests a more task-oriented representation of involvement and solidarity in the Namahn project.

Finally, open greeting only correlated with the involvement marker in the Namahn data ($r = 0.30$), whereas open greeting correlated with involvement ($r = 0.27$), solidarity ($r = 0.27$) and sociability ($r = 0.25$) in the Society data. Thus in the Society communications, open greeting behaviour varies with feelings of involvement, solidarity and sociability, whereas in the Namahn communications, open greeting behaviour only varies with involvement. This difference also suggests a more task-oriented representation in the involvement marker in the Namahn communications.

Thus in conclusion, the fact that sociability, involvement and solidarity all show intercorrelations in the Society data, whereas only involvement and sociability show a (relatively small) correlation in the Namahn data indicates a difference in the balance of task and social representations in the two projects. Dual representation of task and social components in the involvement and solidarity markers explains the difference. Longer emails show higher involvement and solidarity in both projects, but only higher sociability in the Society project. Open greeting behaviour varies with involvement in both projects, but only with solidarity and sociability in the Society project.

Thus the variations in relationships between the communication markers on these projects suggests that in the Namahn project, the task dimension is more dominant in the balance between the task and social dimensions than in the Society project.

7.3.3 Task profiles and project feedback

In section 3.9, I have argued that relative email frequency on networked team projects is representative of relative activity. Figure 7-2 presents the relative activities by purpose for each of the projects. I have removed the "Other" category, as this had a relatively high number of emails related to other Society business, whereas there was only 1 email analysed in the Namahn data which related to other business. Frequency in the "Other" category reflects the data collection method rather than the activities on the project. Secondly, I combined the *Product/SME* and *Review discussion* categories in the Society profile. The SME discussions reviewed course content, but as the handouts represent the course material, review of course content is essentially the same as review of handout content. This point might be debatable, but it does not change the fact that SME discussions are task oriented rather than socio-emotionally oriented. They contribute to the task goal, rather than group maintenance, so having combined the review and SME discussions from the Society project does not invalidate a comparison of the task and group maintenance profiles.

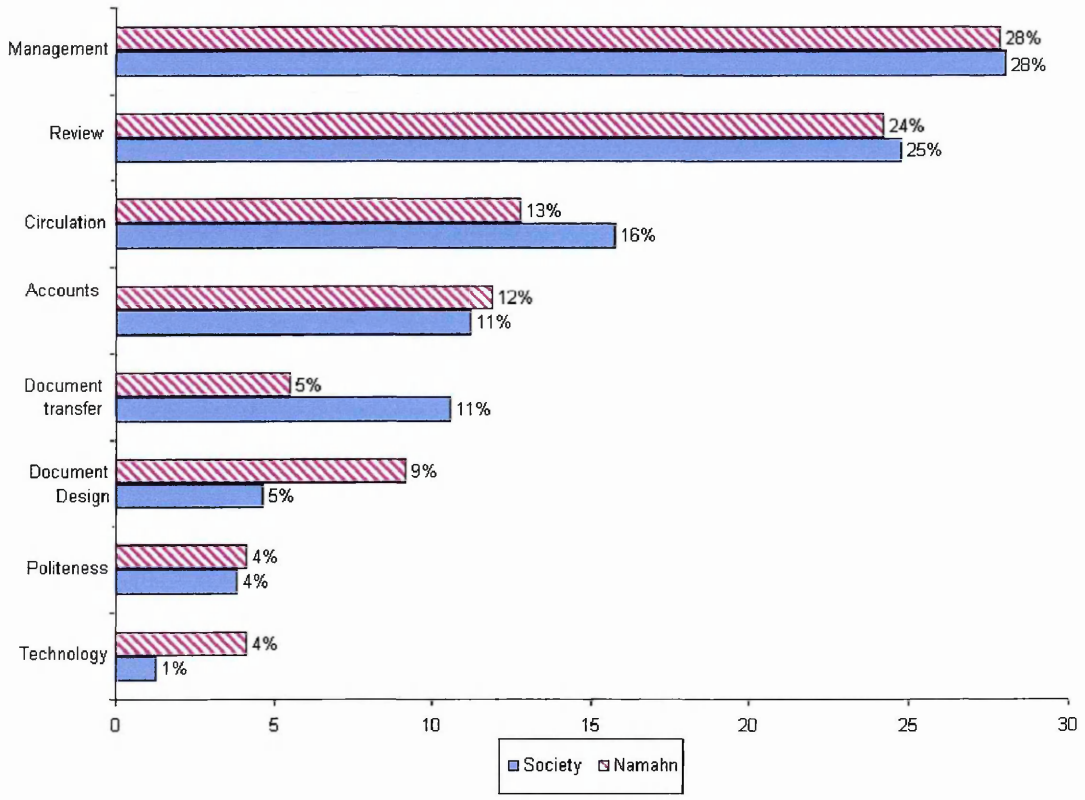

Figure 7-2: Relative activities by purpose on Namahn and Society projects

Of particular interest in Figure 7-2 is the similarity between the task profiles. Activity on group maintenance tasks, *Management* and *Courtesy/Politeness* are identical. *Circulation* and *Document transfer* vary, with higher percentages for the Society project, which is explained by the larger number of team members and the higher number of lecture and workshop handouts to be circulated. *Accounts* representation, interestingly, is almost identical, in spite of the low

rating for the importance of budget in the Society project (appendix K line 98). There was also higher activity on document design in the Namahn project, which reflects the focus of technical writing consultancy on document design. Writing handouts for the Society project was an additional task to developing the course, rather than the focus. Additionally, document design was rated as contributing only average value towards the goal by the Society interviewee (see appendix K line 80-81).

The main conclusion to be drawn from comparing these two profiles of task activities is that the relative effort and time expended on the group maintenance and task dimensions do not differ. The Society project was equally dedicated to accomplishing task oriented activities as the Namahn project, in spite of the evidence that the Society project shows a higher profile of the social dimension in the nature of its communications.

Further, in both the Namahn questionnaire (see appendix EE Section 3) and the Society interview (appendix K line 80-107), I asked respondents to rate different aspects of the projects: how well the project goals were met, how well personal goals were met, how well the budget and deadlines were met and how well the project compared to other similar projects. Responses from project leaders are presented in Figure 7-3. Texts to the right of the bars show the respondents' perception of the value of the measure, i.e. the importance of meeting the deadline or the budget etc. The responses show an overall positive valence in the leader perceptions on the aspects of performance queried, suggesting that both projects were highly successful, i.e. a more highly focused social dimension in the academic project did not detract from overall performance.

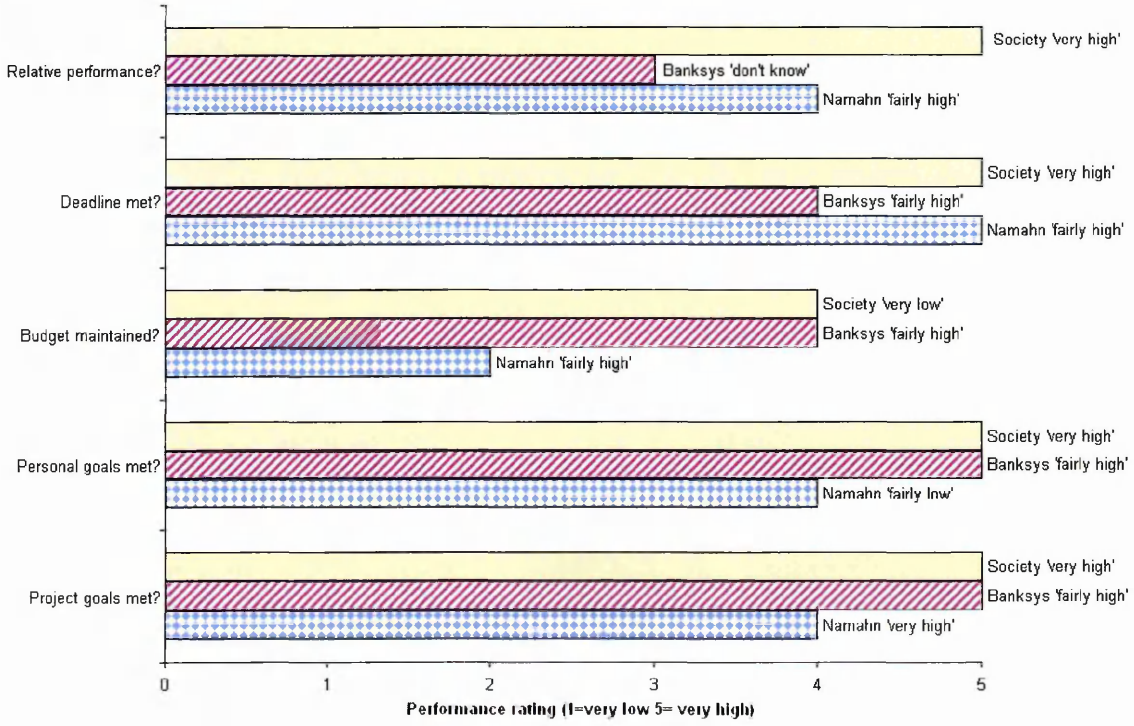

Figure 7-3: Project and Course Leader performance ratings

(texts above columns are perceived importance of this performance measure)

7.3.4 Social-task balance

Overview

In Nystrand's social interactive model of writing, writers anticipate readers' needs, and meaning and interpretation is a shared social reality, the meeting of writer intentions and reader interpretations. In sections 5.4.2 and 6.4.2, I concluded on the balance of the social and task dimensions of the projects based on how team members adapted communication behaviour to task and socially oriented goals. These conclusions were based on how empirical representations of communication behaviour varied with categories of writing influences and are included here to describe the social-task balance on each project.

Commercial project social-task balance

In the commercial project, elaboration, formality and sociability were used to adapt email style to help create a shared understanding between the writer and reader. While sociability was adapted for *Receivers*, one component of the social marker, which contributes to the building and maintenance of relationships, self-disclosure, was missing from the data. Writers did not vary greeting style, involvement or solidarity to achieve a shared understanding with their *Receivers*. There were clear norms of behaviour in open greeting behaviour (see Figure 3-5 to Figure 3-8 in section 3.7.2), showing a social conformity in the team, which may have inhibited individuals from adapting their greetings extensively for an individual *Receiver*. There was also evidence to suggest that functional and organizational factors influence greeting style and involvement. The Kruskal-Wallis tests showed significant differences for both open and close greetings, and involvement, with both *Direction* and *Purpose*. *Purpose* also affected adaptations in solidarity.

Involvement may reflect one or both of the task and social dimensions in the communication. Although writers adapted involvement with the *Audience* size, they also adapted involvement with *Purpose* and *Direction*, suggesting a tendency towards task involvement rather than socio-emotional involvement; involvement did not vary by *Receiver*, suggesting a lack of representation of the relational or socio-emotional element of the marker in this project.

The fact that there were only variations between *Receivers* for elaboration, formality and sociability, rather than for all of the communication markers suggests that the balance between social and task dimensions in the commercial project may lean towards the task dimension, with little socio-emotional communication.

The team members reported a general positive perception of group cohesiveness, although one team member had reservations. Greeting style, involvement and solidarity were not adapted for readers, and involvement, solidarity and sociability marker frequencies were relatively low (see

Table 7-3). I therefore concluded that the social dimension on this project had a relatively low profile compared to the task dimension.

Academic project social-task balance

In the academic project, the team adapted communication behaviour for intended *Receivers* using all the markers I studied, demonstrating the team's practice of social interactive writing behaviour. Involvement varied with both the *Receiver* and the *Purpose*, suggesting both task and socio-emotional involvement in this project.

Writers adapted their email style by *Direction* using all the markers, additionally demonstrating social interaction at an organizational level. This, together with the frequencies of greeting behaviour shown in Figure 3-9, Figure 3-10, and Figure 3-11 suggests the influence of social norms in addition to adaptation by *Sender* and *Receiver* and the *Sender-Receiver* relationship.

Solidarity shown in communications varied in this project with all the writing influences, involvement, solidarity and sociability marker frequencies were relatively high (see Table 7-3), and solidarity was perceived to be high by the interviewee.

I conclude from the high solidarity, high number of marker adaptations with both *Purpose* and *Receiver*, and in particular the adaptation of involvement with both *Purpose* and *Receiver*, that the task and social dimensions of social interactive writing are evenly balanced in the academic project.

Comparing the social-task profiles

Thus based on comparisons of the communication behaviour adaptations in task and social dimensions and solidarity data for the two projects, the commercial project shows representation of more task than socially oriented communication and the academic project shows an even social-task balance.

7.4 Chapter review

7.4.1 Overview

This chapter has drawn together document evaluations and conclusions drawn on the social-task balance of two networked team writing projects to address the hypothesis:

> *H3 = Social interactive adaptations and pro-social behaviour in a writing team's emails are reflected in the relational metadiscourse and social desirability of the document produced by the team.*

The commercial project produced Owner and Service manuals for the installation of a payment terminal in shops, and the academic project produced the handouts for a postgraduate clinical training course.

7.4.2 Document evaluations

Critical evaluations of the documents against Sless's social desirability attributes highlighted strengths and weaknesses. Both documents are attractive, physically appropriate and highly productive, although there is some scope for improvement in both to support ease of use and efficiency, for example by improving navigation with section numbers. The strongest attributes of the course handouts are their social appropriateness and respect for the reader; the weakest attributes are credibility and efficiency. The manuals have strong corporate credibility, with some scope for improvement on social appropriateness and respect for the reader. The organization has a higher profile in the manuals than in the course handouts, and together with the high strengths of the course handouts in terms of social appropriateness and respect for the reader, I conclude that the academic documents show a more even balance than the commercial documents, between the task and relational elements communicated to readers.

7.4.3 Social-task balance

The significantly higher mean marker frequencies for involvement, solidarity and sociability suggest that these feelings were represented more in communications on the academic project than in the commercial project. Despite the higher socio-emotional representation on the academic project, relative effort and time expended on the social and task dimensions did not differ between the two projects and feedback from the team leaders suggested that both projects were perceived as being highly successful. Variations in correlations between the communication markers also suggest a task orientation in the commercial project.

Analyses of communication markers to profile adaptations with writing influences also showed differences between the projects. In the commercial project there were only variations between *Receivers* for elaboration, formality and sociability. While sociability was adapted for *Receivers*, one component of the social marker, which contributes to the building and maintenance of relationships, self-disclosure, was missing from the data. Greeting style, involvement and solidarity were not adapted for readers. Although writers adapted involvement with the *Audience* size, they also adapted involvement with *Purpose* and *Direction*, suggesting a tendency towards task involvement rather than socio-emotional involvement; involvement did not vary by *Receiver*, suggesting a lack of representation of the relational or socio-emotional element of the marker in this project. The team members reported a general positive perception of group cohesiveness, although one team member had reservations. Overall, therefore, the social dimension on this project appeared to have a relatively low profile compared to the task dimension.

In the academic project, writers adapted their communication behaviour for their intended *Receivers* using all the markers I studied, demonstrating the team's practice of social interactive writing behaviour. Involvement varied with both the *Receiver* and the *Purpose*, suggesting both task and socio-emotional involvement. Solidarity shown in communications varied with all the

writing influences and was perceived to be high by the interviewee. The task and social dimensions on this project thus appear to be evenly balanced.

7.4.4 Answering the research question

This chapter has explored the concept that social interactive adaptations and pro-social behaviour in a team's emails are reflected in the relational metadiscourse and social desirability of the document produced by the team. A social dimension in team working contributes positively to the outcome (see section 2.3.2), so that the social-task balance on a project also affects the team performance. Here I draw together conclusions on the social-task balance and social desirability of the documents produced by the two projects.

Mean frequencies for socio-emotional communication markers were higher in the academic project, and correlations between the markers suggested more socio-emotional marker components than in the commercial project. Comparisons of adaptations of socio-emotional communication behaviour with receivers and purposes, and feedback on solidarity in the projects suggested a more highly profiled social dimension in the academic project. Finally comparing the task profiles of the projects did not suggest that the task focus was lower in the academic project than in the commercial project. I conclude therefore that the academic project demonstrates both task and social profiles evenly and that the commercial project leans towards the task dimension.

Both projects were successful, although both documents had scope for improvement. High strengths of the course handouts in terms of social appropriateness and respect for the reader led me to conclude that the academic documents showed a more even balance than the commercial documents, between the task and relational elements communicated to readers.

Thus the social-task balance on these projects, described by the social interactive adaptations and pro-social behaviour in team emails, were reflected in the relational metadiscourse and social desirability of the final documents. Social interactive adaptations demonstrated in socio-emotional behaviour in the emails of the academic project were also demonstrated in the final document.

The evidence here is drawn from empirical and qualitative evaluations from two case studies. Combining the indicators and comparing the data for these two projects shows that social behaviour on networked team writing projects, identified from email communication behaviour, is reflected in the final document. In this case, the project which showed a higher socio-emotional content and social interactivity in communications also showed higher social appropriateness and respect for the reader in its final documentation. To prove this hypothesis scientifically and quantitatively requires repetition on multiple writing projects, using reader-focused, rather than expert-based document evaluations by a single researcher. The potential which socio-emotional communication can contribute to project performance is discussed further with examples in section 8.3.

8. Discussion

8.1 Chapter overview

The research question I addressed with this work was "Can we learn about the influence of team culture on virtual team writing from content analysis of email communications during projects?" In this chapter I draw together some observations from my work to answer this question. First in section 8.2, I discuss some aspects of developing the email analysis tool, highlighting areas for future research. In section 8.3 I discuss the social-task balance concept at a project, individual and email level, from which I develop my own personal recommendation for professional email communication. Two main aspects of my work have been profiling social dimensions from email behaviour in academic and commercial contexts, and the concept that social-interactive adaptations in email writing are reflected in the final document; these topics are discussed in sections 8.4 and 8.5.

8.2 Developing the email analysis tool

8.2.1 Developing the formality score

Interrater agreement on components of the formality score in the commercial project was inadequate and attributed to unclear representation of automated signatures. This component was not used in the improved formality score for the academic project. Interestingly, however, interviewee ratings of components (appendix K line 188-189) varied slightly from the values assigned on the continuum from spoken to written style in the academic project, suggesting that there is scope for improvement to the design of the score. There are two foci worthy of attention in future research and these are the holistic interpretation of greeting and signature markers and the influence of time factors on formality, and I explain both these concepts here.

The formality design I have used assumes a stable interpretation of formality for greeting and signature components. It is particularly interesting, however, that raters in the commercial project showed variations in their interpretation of formality for automated signatures according to whether a manual signature was also present. It may be that the salutation styles cannot be decontextualised from each other, and that their meaning is created by the overall tone which they present combined. To measure this accurately would require a more holistic scoring of formality for different combinations of components. An experimental design using emails with varying combinations of open, close, manual and automated signatures, and asking readers to assign single formality scores for the different combinations could be used to explore this issue.

The second area for future research relates to distinguishing between the effects of different time scales. There are three time scales relevant to the study of email style, the immediacy timeline (how quickly the correspondents turn-take), the socialisation timeline, and the evolution timeline of email style as an emergent register.

In the commercial project, there were examples where the formality style altered according to the rate of message exchange. In communications which were exchanged more rapidly, the style became more conversational. This immediacy, which Daft and Lengel (1984) specify as one of the media richness attributes, fluctuates *within* a media-type in professional practice, thus introducing another variable which needs to be taken into account. This first example shows the usual style in emails from a subcontractor in the commercial study:

Text 3975, 2 December 2003 Dag XXX, Hierbij de eerste twee .... Met vriendelijke groeten, [Full name] Full title] Full automated signature]

In the fourth exchange between the same two individuals on the following day, the same author uses no greetings whatsoever, and launches straight into the message with a more conversational style, i.e. no open or close greetings and no signatures:

Text 3783, 3 December 2003 Het is een probleem dat wel vaker voorkomt bij Word....

The rate of email exchange thus influenced the change from a formal written business style to a more conversational style.

Time has other influences on formality. Danet, in her analysis of two-way interpersonal exchanges between herself and a correspondent demonstrated a change in formality over time (Danet 2001a p84). In both the academic and commercial case studies in my research formality was shown empirically to vary with socialisation phase, and I present a qualitative example here. The same author in the examples above, who normally included his title and automated signature without the immediacy of exchanges illustrated above, finally dropped his formal pre-coded signature, when responding towards the end of the project to compliments about his work. Thus the formality level changed again:

Text 3707 15 January 2004
Dag XXX,
Dank u voor het compliment; het was dan ook een prettige samenwerking....
Groeten en tot ziens
[First name only]

I've used these examples to illustrate the complexity of formality; they also serve to illustrate the iterative nature of discourse and social behaviour. Formality is derived from a complex balance of subtle adaptations in behaviour, such as finally overcoming organizational norms and dropping an automated signature.

The third time scale relates to the evolution of email style generally. Danet concludes that email style is in a state of transition (Danet 2001a p93) and Crystal also agrees that "The evolution of email style is in its infancy" (Crystal 2001 p127). This research has therefore only glimpsed a snapshot on the changing backdrop of an emerging register. As such, it can only inform from observations of practice as it is today.

All of these time scales and their interactions are influential on email style and need to be taken into consideration in email research.

8.2.2 Interdependencies and dual representations

Hyland (1998 p230) discusses the polypragmatic character of discourse and warns that "Language use is 'fuzzy", which means that meta and propositional discourse cannot always be distinguished and that types of metadiscourse do not exclusively perform either textual or interpersonal functions" (Hyland 1998 p230). Differentiating between representation of social and task dimensions was also fuzzy in my research.

Quantitative and qualitative data from the commercial case study suggested that the relationship between the social (or group maintenance) and task dimensions of the project is complex and cannot necessarily be analysed in a discrete way. For example, an interviewee described reasons for expressing sociability and formality for particular *Purposes* in the project, thus using socio-emotional characteristics of communication for elements interpretively coded on the task dimension. Writers were more formal in communications on tasks completed early on e.g. *Accounts*, when they didn't know each other so well, and for purposes over which they felt more anxious, e.g. *Review.* Task and social contributions at a team level are thus interdependent and cannot be analysed discretely, but rather need to be considered in a holistic way.

There were also other types of dual representation which repeatedly emerged while interpreting the data from both case studies. These included where a single marker had subcomponents,

where a single marker represented both task and social elements or where a marker had some confounding effect, and I explain these points here.

A positive correlation on the academic project and negative correlation on the commercial project between formality and sociability suggests that individuals in the academic project used social building strategies to develop relationships, whereas in the commercial project, social building units represented the closer less formal relationships.

Social building units included general markers (such as "Happy New Year"), apology, courtesy, humour and self-disclosure markers. Studying the variations in use of these markers with time in the commercial case study showed different trends. Politeness/courtesy markers increased with socialisation whereas use of the more general social marker decreased. An email author in the academic project rated high on sociability markers, but was rated low on sociability by the interviewee. Studying this individual's emails in depth showed that the type of sociability markers represented were almost entirely politeness/courtesy markers. The markers contributed to a positive tone, thus helping to build and maintain relationships, but the email author's style did not encourage the interviewee to rate him as particularly sociable. The only authors who were rated as sociable in terms of social markers and also rated sociable by the interviewee were those who had included self-disclosure. Different social markers thus represent slightly different strategies in pro-social behaviour. Self-disclosure is influenced more by individuality and the *Sender-Receiver* relationship, whereas courtesy is influenced more by social norms of behaviour. Further research is required to find ways of systematically distinguishing between the underpinning pro-social strategies in uses of the social building markers used in this research.

Other markers showed similar complexities for interpretations. Greeting length may represent both formality and value and effort attributed to communications, and varies with influences on behaviour such as *Direction* and *Audience* size, suggesting the influence of social norms as well as the Sender-Receiver relationship. Clear norms of behaviour were apparent on both projects from profiles of the frequencies of greetings and signatures. However, open and close greeting behaviour did not adapt consistently in the academic case study. Variations in style of open and close greeting highlight that the markers comprise different socio-emotional components. Close greeting length did not vary significantly between the commercial and academic projects, suggesting that it is subject to more pervasive norms. Interestingly, however, close greeting length correlated positively with involvement, solidarity and sociability in both the academic and commercial projects.

Solidarity and involvement marker representations may be confused by the possibility of only including one or the other, and solidarity markers may be used as formal representations of an entity rather than a representation of solidarity. Additionally, solidarity markers may be used as a strategy to build solidarity or to maintain solidarity, and finally, involvement markers represent both task and socio-emotional involvement in an individual's contribution, but not the extent of the contribution in the project.

These complexities of the markers and their interpretations mean that they cannot be interpreted in isolation, or in a purely quantitative way. Rather they need to be searched in a holistic way, because each marker comprises slightly different and sometimes multiple representations, which contribute to the whole picture of an individual's style of email communication behaviour.

8.2.3 Face to face confounding effects

One of the important factors in this research is that I was dealing with the real world, which paradoxically meant that the working environment I studied was a mix of the real and the virtual.

The projects studied were not exclusively virtual, an artificial scenario often used in CMC research (for example in experimental research by Walther, Burke and Chidambaram etc. see section 2.3.4). The value of FtF meetings in professional work is recognized as beneficial to professional projects (Walther and Parks 2002 p556). The commercial project had five FtF meetings in total, three between the technical writer and client, one between the project leader and illustrator, and one between the illustrator and client. The academic project had on average two FtF meetings per year.

This aspect of real world researching had some confusing implications. For example, the academic interviewee perceptions sometimes varied from the email data interpretations, either due to perceptions being altered by FtF meetings, or to communication behaviour being altered by FtF meetings. I explain examples of both possibilities here.

The academic interviewee perceived an Academic Faculty member (AF2) as highly active on the project, although this person was only responsible for 2% of the email communications. When asked to comment on this, the interviewee explained that this was the person with whom she had the most FtF contact. Email data also showed that formality and sociability were low between the Course Leader and Administrator, who also had regular FtF meetings. Social building units were therefore unnecessary in the email communications between these individuals, because they were able to build their relationship in FtF meetings. Formality and social building units were also low in exchanges between the Course Leader and a co-located colleague, which the interviewee attributed to their working together in the same location. Influences of FtF meeting on CMC behaviour were also apparent in adaptations which coincided with FtF meeting dates (discussed later in section 8.4).

These examples highlight that FtF contact needs to be taken into account when interpreting social markers in emails, bearing in mind that the relationship is not solely dependent on relational communication in CMC. Additionally team member perceptions will not necessarily match interpretations from CMC behaviour, as they will be coloured by experiences of FtF behaviour. Rooksby argues that textual style will never be wholly disclosive of self, because it is only a subset of all styled performances, and becomes the only part of an individual's styled performances visible through CMC. She argues, therefore, for attention to this reduction, the "transformative effects of textual style on textually-mediated social relations" (Rooksby 2002 p10),

suggesting that "attention to these transformations may suggest new ways of reading and writing, and ways to avoid the interpretive pitfalls characteristic of textual communication" (Rooksby 2002 p10). Differences between perceptions of FtF social relations and perceptions of textually mediated social relations is an interesting area of study in itself. Another area for further study is the ratio of FtF meeting and CMC in projects and its impact on both the socio-emotional needs in CMC and project performance. Variations in CMC behaviour with increasing FtF contact and the effects on performance of varying combinations of CMC and FtF communication need to be researched.

A final point of relevance to FtF meeting is that the academic document studied in this research was designed for use in a FtF meeting between the author and reader. This writing influence, the combined FtF and textual delivery media, may have influenced the social interactive adaptations used to contribute to the document's social appropriateness and also raises another interesting area for writing research. Social interactive writing behaviour may be applied more in scenarios where writers anticipate FtF meetings with their readers.

8.2.4 Language and culture

Communications were in Dutch and English in both the commercial and academic contexts. In the commercial project, English emails were written by Dutch and English speakers. In the academic project English emails were written by individuals of various nationalities.

Based on Brown and Levinson's work (1987), I did not expect differences in pro-social strategies between languages. Some differences did emerge, however. For example, an American team member showed relatively high involvement and Dutch speakers showed relatively high solidarity. Research by Nickerson (2000) has shown relative higher exclusive use of the marker "we" by Dutch authors, which qualitative study of email examples in the academic study suggested was also the case in this research.

Additionally, findings related to language were not consistent across the two projects. Elaboration, solidarity, sociability and formality were higher in Dutch emails in the academic project; sociability and formality measured by open greeting length were higher in English emails in the commercial project. Thus the results show higher pro-social behaviour in CMC of Dutch writers than English writers in the academic project, but higher pro-social behaviour in CMC of English writers than Dutch writers in the commercial project. This concept needs exploring further in multiple commercial and academic contexts to establish whether there is a language dependent difference in pro-social CMC behaviour between the two types of discourse community.

8.3 Social-task balance: academic and commercial contexts

The social dimension in team work contributes positively to performance (see section 2.3.2), so that the balance between social and task dimensions on a project also affects performance. In this research I have focused on the relational metadiscourse and social desirability of

documents as outcome measures. Social and task dimensions also describe official and personal email styles. Danet (2001a) expects both official and personal emails to become more speech-like and the distinction between the two letter types to become less sharp. Baron predicts the likelihood of two styles of email, one formal (edited) and one informal (unedited). She points out that frequent email users may switch off automated editors and may even choose not to manually edit, thus communicating in an informal way, whereas on the other hand "a contract is still a contract" (Baron 2001 p242), requiring accuracy, editing and thereby more formal communication.

This research has shown a more task focused approach to email interaction in the commercial setting. Nickerson (2000) found a "certain amount of relational or non-propositional content, intended to maintain the social system within the corporation, i.e. the patterns of corporate social relations between employees" (Nickerson 2000 p153). Interestingly, also, work on organizational email as early as 1986 by Sproull and Kiesler showed a ratio of work to non-work related emails of 6:4 (n=1248). Non-work related emails covered personal topics such as recipes, advice on where to get a second mortgage etc. (Sproull and Kiesler 1986). This finding may throw light on why emails in my research appear to have task-oriented foci in the commercial setting and has more to do with the email data collection than the absence or presence of relational communication. It may just be that socio-emotional content, requiring a different style to work-related (more official) content, requires separate communications, and that writers find it difficult to combine different styles for different purposes within a single communication.

However, there is also evidence from the literature to suggest that email style differs between academic and business contexts. Gains (1999) showed a distinction between academic and commercial discourse communities. In a small scale study of 116 emails in academic and commercial settings he found that academic emails were less formal with more social chat, i.e. more like conversation, whereas emails from a commercial environment were more like written business language. He comments that the academics in his study treated emails more as "throw-away" communications (Gains 1999 p99). Geisler et al. (2001), however, point out the permanency of emails and how they create an organizational memory: "What is recorded in such documents as minutes becomes the official understanding of what has happened or what will happen so that texts are used to shape members' understanding of the organization and its past and future activities" (Geisler et al. 2001 p279). She points out that such records may even be used in disciplinary action. This brings me back to Baron's comment "a contract is still a contract". This permanency of email over conversation may inhibit team members in business contexts from including relational content.

In this research, it was remarkably apparent that self disclosure, involvement and solidarity were more frequently represented in emails written in the academic than in the commercial project. The merging of personal and professional styles which Danet (2001a) has predicted is already

apparent in communications from the academic context in this research, and I provide an example here:

> Text 4895
>
> Hi ZZZ,
>
> It is good to see all this work being done. I am a bit overwhelmed at the moment with moving details and paperwork....
>
> Today we brought the boxes with the stuff to take to the [location]. So I spent the last 4 days packing boxes, selecting clothes etc. but that is now over and done...
>
> The visit to [location] was a bit strange in a way that it is different to meet with a group you have not actually worked with before. I think that after the first initial hesitations we agreed on a good workplan. I have to say if only half of the work can be performed the whole trip will be more than worth the effort. Also from a family point of view, we managed to find a house in a very family friendly neighbourhood with a school at the other side of the street. [Spouse] was with me on the trip and is convinced that the location will be ideal for the children.
>
> Your suggestions concerning the clinical exams are great and would be very helpful for me. For the handouts I would only need a description of the patient history and a filled out clinical exam form, conforming to the forms used last year as well as the video observation sheet (I will forward you a form in attach). I need to deliver the handout material by [date] to [course organizer for 2004]. So I would need to receive this as well as the revised text of the handout text of your presentation by [date]. Would that be possible? The video is for the presentation and that can be delayed a bit longer. You can best just send them to the [author's new address].
>
> I think that the [clinical topic] and [clinical topic] would ideally be average [data type] in excel so that I could put them in the same format as agreed for the handouts, last year. Would that be possible?
>
> Thanks again for all you help with the course. I will try to update some handout notes (esp the patient cases of last year) and I will circulate these for your comments beginning of next week.
>
> If you need my brains for productive and inspiring ideas....I have to check if I accidentally packed them. Feels like that.
>
> [First name]

This is a typical example of the mix of personal and task-oriented content in communications on the academic project. Notice in this example that the author twice asks for confirmation that the reader will be able to meet her requests ("Would that be possible?"). The personal content of the email thus does not detract from the focus on the task, but adds a relational component to the communication. This example demonstrates how the empirical descriptions of relational

communication in the academic project can be substantiated by closer inspection of a qualitative nature. The numbers showed higher means for social markers in this project (see section 7.3.1) although the task profiles were not dissimilar between the commercial and academic projects (see section 7.3.3).

In contrast, I present here two examples from the Keyware writing project reported in chapter 4.

Text 3542

Dear XXX,

Following [Sales Support Manager]'s review and your own comments, I've revised the manual, and attach an updated copy.

Minor corrections were made throughout the document following our conversations, but the main changes are to Section 6 How to Administrate the [Product]. I've removed all the instructions and substituted "Helpful Questions and Answers" in an attempt to pitch the content more appropriately for a System Administrator, and removed duplicated content.

Warnings and important information which were included in the instructions have been included with the Field Descriptions.

Please let me know what you think.

Thanks,

K

Text 3544

Kirstie,
For the administration interface I don't think html is the appropriate format. I will come back on this later today.
[First name]

In the first example above, from me to my manager, I show relative informality in my signature and some courtesy in my closing, but there is no true relational content. With the task focus and formality of the main content, it would be quite difficult to incorporate any relational content. Changing to a more conversational style would allow the opportunity to include a query or personal comment not directly related to the purpose of the email and thus contributing towards building the social dimension of the team. The second example is a typical email from my manager to me. Between us we appear to have fallen into a communicative behaviour which excluded the possibility of informal exchanges. This behaviour may have been governed by our individual attributes, the manager-subordinate relationship, the organizational culture or other causes, but resulted in a functionally efficient style, which lacked socio-emotional content, an element known to benefit team performance. The first email example given above was written by a female on an academic project. The second was written by a female on a commercial

project, and the third by a male on the same commercial project. Future research also needs to address the influence of gender on team social interactive behaviour.

In the commercial study (chapter 5), empirical data extracted from email content analysis suggested fewer social interactive adaptations, and the frequencies of social markers were lower than in the academic project. There were however, some good indications of sensitivities to emotional elements in the commercial project. For example, empathy was shown when a subcontractor experienced difficulties, reservations regarding group cohesion were expressed by a team member, and feelings of anxiety over review were voiced.

The overall differences between the commercial and academic projects may not be underpinned by differences in the openness or communication behaviours of the team members (Slater's "ideas" or "best liked" person; 1965; see section 2.3.2), or even social norms in the client-supplier relationship, but rather a tendency in business settings to separate the task from the social in written communications, to focus on functional efficiency. The task focus of projects may also be nothing to do with deficit theories of mediated communication, which suggest an inherent paucity of socio-emotional communication in leaner media and hypercommunication which focuses on task. Rather it has more to do with the existing social-task balance on the project.

Although this task focus may be deemed the optimal approach in business settings, there may be value in adopting a more conversational style in written professional emails, and I illustrate this with an example here. In the commercial study, a review meeting was suboptimal. In spite of a very successful project, there was also some disappointment in the terminology and this was reported in the final debriefing. Informal exchanges by email between the project leader and the copywriter, and also between the project leader and the client reviewers may have offered the opportunity to exchange ideas on terminology and correct this earlier in the project. An equivalent scenario from the work of Kraut and colleagues (1988; see section 2.3.3) would be the informal exchanges in the corridor between co-located team members. These authors reported that researchers they interviewed supervised subordinates and coordinated with peers during casual hallway and lunchroom conversations, just as often as during formal scheduled meetings (Kraut et al. 1988).

The very nature of what Danet calls "written conversation" (Danet 2001 p 57) in email may provide the informal unplanned communications afforded in co-located teams, which Kraut and his colleagues (1988) concluded is so important to the team's welfare. The Society author in the above example was prompted by task-oriented motivations to contact the addressee, but took the opportunity to write about personal issues. Opportunizing on the task-motivated exchange to build the social and task dimensions contributes to Walther's social information processing theory in long term projects.

Finally, however, in recognizing that informal unplanned communications may benefit the team, we must also remember that performance increases with team cohesion and then decreases (see section 2.3.2). Researchers have argued that the task focus of email communication is the

very benefit which improves performance, as I have discussed in section 2.3.4. Both the documents evaluated in this research were extremely professional, albeit with some potential for improvement. There was a difference in the relational approach to communication in the final documents, which paralleled the relative relational communication adaptations and content in the projects. Whether this trend can be shown in multiple projects requires further research. Also of particular value to test recommendations from this research, would be a study in which different corporate cultures allowing and disallowing personal email communications were researched and project performance evaluated. My hypothesis would be that the organizations whose cultures allow and encourage relational communications would show higher performance.

Damrau and many others recognize the value of socio-emotional communication in virtual teams (Nickerson 2000 p188; Panteli 2004 p78; Pauleen and Yoong 2001 p197).

> *Team socializing is difficult for virtual teams…. A few minutes of sociable conversation helps team members feel connected and learn their colleagues' personalities, which can, in turn, lubricate their online communication with each other (Damrau 2006 p13).*

Walther argues that the deficit theories, social presence, cuelessness and media richness theories, assert "that the structure of the medium alters the nature and interpretation of messages, it implies that such effects are inherent and constant whenever people communicate using computers" (Walther 1995 p188).

> *The effects of the CMC channel depend not on bandwidth alone, but on the interactions of media characteristics with social contexts, relational goals, salient norms and temporal frames that promote or inhibit the strategic use of CMC in relational supportive or detrimental ways (Walther and Parks 2002 p556).*

My research supports Walther's criticism of deficit theories; writers are adapting their email style with many different factors and demonstrate socio-interactive writing behaviour either at an organizational or individual level. The fact that this research has shown many adaptations in communication behaviour with writing influences supports adaptive structuration theory. Writers are not restricted by the medium in their communication behaviour, but adapt with the technology available to them.

However, there is no opportunity for informal unplanned task-checking, for example to check whether a particular term being used in a document is technically correct and socially appropriate. This lack of informal opportunity combined with social norms of behaviour, inhibit our opportunizing on the benefits of email. Email is written conversation. Email shows a "versatility of discourse styles" (Yates and Orlikowski 1993 p13); this characteristic of email offers opportunities otherwise missed in non-mediated communication; this is the area for exploration in the training of CMC. Social and task elements are combined in teams, in individuals and within single exchanges without losing focus on the task. Social dimensions contribute to performance and thus benefit the project. Professional email communications can and must offer more than a traditional business letter. They provide the social construction of the team and must therefore provide both social and task dimensions and the opportunities for

informal unplanned communication without inhibition or censure. Walther points out how the asynchronous nature of email supports this: "...temporal commitments become discretionary, and task versus interpersonal interaction becomes ... de-regulated; both task and social exchange may exist without one constraining the time available for the other" (Walther 1995 p199).

The balance between the task and social dimensions of team work is defined by the combined task and social representations in the emails on the project. At a team level, from Slater's work (1965: see also section 2.3.2), it seems that the combined contributions of the task expert and the sociable expert affect the team's well-being. Having individuals who demonstrate not only task but also socio-emotional communicative skills is also a benefit to the project. Burgoon and Hale point out that "A person who is very task oriented may still demonstrate sociable tendencies" (Burgoon and Hale 1987 p40), which brings me to my recommendation for professional email communications in semi- or virtual team writing projects.

My recommendation from what I have learnt from this research is for professionals working by email to veer towards conversational style in written communications, towards a style which is more likely to prompt unplanned informal socio-emotional content and help build and maintain relationships in the team. Such an approach helps to open doors and overcome barriers which individuals feel when emailing professional contacts from behind the façade of tradition or organizational norms. Exchanges characterised by such openness are more likely to afford more frequent informal exchanges, providing the opportunities for checking minor details and exchanging ideas, thus contributing to the quality of both task and social dimensions.

8.4 Micro-analyses: academic and commercial contexts

Contextual information gathered from team members in pre- and post-analysis interviews, and combining evidence from multiple markers helped to interpret the detailed adaptations in communication behaviour on the projects. Social influences on behaviour were validated by participant feedback in both studies. Participants justified changes in communication behaviour following changes in team composition, FtF meetings and milestones in the projects. Using member-checking where team members are shown analyses and significant variations in behaviours to develop interpretations thus has potential for identifying and correcting problem areas in long term writing projects

Across context comparisons at this detailed level, may also identify social norms linked to email as a genre rather than to a particular context of work or discourse community. For example, in this study, results from both the commercial and academic projects showed a decrease in involvement and social building units with increased audience size. Both projects also showed the same unusual trend in solidarity markers, which were higher with an audience than without an audience, but lower in emails addressed to more than one individual.

Examples of interpretations which were reinforced by participant perceptions reflected the power hierarchy of the business relationship. In the commercial project, in terms of elaboration, the supplier made more effort and attributed more value for communications to the client, than for communications in house, and the client made less effort and attributed less value to the communications for the supplier than for communications in house. Profiles for elaboration by relational direction thus reflected the power hierarchy of the business relationship. The equalization effects predicted by Sproull and Kiesler's (1986) lack of social context cues hypothesis are therefore not pervasive across email communication contexts, possibly attributable in this case to the semi-virtual nature of the project.

There was also a relational difference in the tendency for higher formality in open and close greetings for the supplier to client emails than for the reverse emails, and higher formality in supplier open greetings than client open greetings. Involvement was higher on the supplier side than the client side, describing the supplier-client business relationship, which defines the supplier as the most active on the project. In house (as opposed to external) supplier interpersonal communications, where relationships are likely to be more familiar, also showed more involvement and less formal close greetings.

Also in the commercial data, although writers appeared not to adapt greetings to *Receivers*, greetings did vary according to the writer's organizational context, and the organizational context of the *Receiver*, suggesting that the team was practising social interactive communication behaviour, albeit not at an interpersonal level, rather at the organizational level. Thus the contextual relationship between writer and reader and organizational norms clearly influenced communication behaviour and email style in the commercial project.

Similar organizational communication behaviour patterns emerged in the Keyware data (chapter 4). Emails *written* by the lowest level members were most formal, emails *received* by the lowest members were least formal, and emails sent upwards were most formal. Analysis of these markers thus informs on the communication and social norms, which in this case once again reflect the power relationship in the organizational hierarchy.

Also in the Keyware data, highest level members showed the highest involvement and the two middle levels showed relatively higher solidarity than the two extreme levels. This observation might be worthy of study to assess what level of involvement at managerial level predicts good team performance. Falling below a certain involvement level might flag a potential problem. This kind of email diagnostics could be used to avoid problems in long term projects.

Social aspects identified from the academic project profiled administrative versus subject matter expertise functions of emails, and the fact that there were two separate organizations involved, the Society and the Conference Organizers.

The methodology also supports the study of time-based adaptations. In this research, I chose to use sequential email communications as a continuum of socialisation believing it to be more relevant than a time line in semi-virtual projects with few FtF meetings. In section 2.3.4 under

"Theories and research", I discussed social information processing theory, which predicts that socio-emotional communication in longitudinal projects will reach the same levels of as FtF projects with time. Walther's work has shown that when CMC participants are interdependent over time, they adopt more intimate and sociable relational behaviour from the beginning of projects and throughout (Walther 1995 p198). Social information processing could therefore explain the difference between the two projects in terms of relative socio-emotional marker frequencies. The commercial project was for a finite number of months, whereas the academic project has been ongoing for a number of years and still is. Thus the long term interdependencies in the academic project, may have improved the relational content in the communications. Apart from the overall higher frequency of social markers in communications on this project, involvement on the academic project showed an overall increase with socialisation phase.

As the academic project is ongoing, it is difficult to map Gersick's (1988) punctuated equilibrium theory onto the socialisation phase. The interviewee reported renewed task goal orientation and increased sociability after the 2002 conference at a point in time between the third and fourth socialisation phases (out of six phases analysed). Phase four showed an increase in effort and value attributed to communications, and increases in solidarity and sociability. Thus FtF meeting and accomplishment of the course and conference in 2002 improved the task orientation and sociability of the project, and these improvements were visible from the communication markers.

Interestingly, adaptations in some of the communication markers on the commercial project mapped two critical periods when groups are more open to influence according to Gersick's theory (1988); the initial meeting and the transition point. Effort and value attributed to the email and solidarity show peaks at the beginning, middle and end of the project. These peaks coincide with calendar dates for initial, midpoint and end FtF meetings held on the project. This influence of FtF meeting on CMC behaviour introduces an interesting area for further research, discussed earlier in section 8.2.3.

Thus meaningful interpretations were developed from situated knowledge and the email data. To determine whether such interpretations at this micro level are consistent for similar variations in different projects requires building a database of trends interpreted from similar detailed analyses on multiple projects.

8.5 Social interactive behaviour and the social-task balance

In this research I have explored the concept that content analysis of emails can describe social interactive team behaviour, which is reflected in the social desirability of the final document. Based on the premise that socio-emotional communication is equally important to team performance as communications over task goals, I used pro-social markers from emails to interpret social-interactive adaptations in metadiscourse with both task and social writing influences.

In Nystrand's social interactive model of writing (1989), writers anticipate readers' needs, and meaning and interpretation is a shared social reality, the meeting of writer intentions and reader interpretations. Textuality of email means that individuals must use their writing skills to communicate effectively for task and socio-emotional objectives. Pro-social markers in emails profile the adaptations in metadiscourse practised in social-interactive writing. Hyland describes metadiscourse as "an important means of supporting the writer's position and building writer-reader relationships". Interpersonal metadiscourse comprises elements of writing which "reveal the author's attitude towards both the propositional information and the readers" (Hyland 1998 p229) and "demonstrates that the writer has taken the prospective reader into consideration" (Hyland 1998 p238). "It is essentially evaluative and relates to the level of personality, or tenor, of the discourse, influencing such matters as the author's intimacy or remoteness, expression of attitude, commitment to propositions, and degree of reader involvement" (Hyland 1998 p229). The way teams adapt email style provides a means to evaluate how they build and maintain relationships, contributing to group maintenance and in turn affecting performance.

To explore whether pro-social behaviour profiled from team email communications is reflected in the final team document, I developed a relatively unintrusive methodology for researching professional writing, an email analysis tool. Using this tool to analyse communications in a networked team writing project (chapter 4), I have shown that communication behaviours vary by *Sender*, *Receiver* and other writing influences, conforming to Nystrand's social interactive model of written communication (1989). This substantiates the study of these markers at a macro or group level to compare the social dimensions of projects, and use them as predictors of team performance.

I have analysed and compared commercial and academic writing projects using this technique, to illustrate the concept of researching social interactive writing adaptations in emails and the potential this method may have to predict social interactive writing in the final documents written by networked teams.

In the commercial project, writers did not vary greeting style, involvement or solidarity to achieve a shared understanding with their *Receivers*. Although writers adapted involvement with the *Audience* size, they also adapted involvement with *Purpose* and *Direction*, suggesting a tendency towards task involvement rather than socio-emotional involvement; involvement did not vary by *Receiver*, suggesting a lack of representation of the relational or socio-emotional element of the marker in this project. Although writers did not adapt greetings for *Receivers*, greetings did vary according to the writer's organizational context, and the organizational context of the *Receiver*, suggesting the writers were practising social interactive communication behaviour at the organizational level.

In the academic project, writers adapted their communication behaviour for their intended *Receivers* using all the markers I studied, demonstrating the team's practice of social interactive writing behaviour. Involvement varied with both the *Receiver* and the *Purpose*, suggesting both task and socio-emotional involvement in this project. Writers adapted their email style by

Direction using all the markers, additionally demonstrating social interaction at an organizational level.

Mean frequencies for socio-emotional communication markers were higher in the academic project, and correlations between the markers suggested more socio-emotional marker components than in the commercial project. Of particular note in the commercial project was the absence of self-disclosure in communications. Feedback on solidarity in the projects also suggested a more highly profiled social dimension in the academic project. Finally comparing the task profiles of the projects did not suggest that the task focus was lower in the academic project than in the commercial project. I concluded therefore that the academic project demonstrates both task and social profiles evenly and that the commercial project leans towards the task dimension.

Feedback from key team members shows that the overall perception of performance was high in both projects. Evaluations based on Sless's social desirability criteria showed a more even balance between the task and relational elements communicated to readers in the academic document than in the commercial document. In this research, the project which showed a higher socio-emotional content and social interactivity in email communications also showed higher social appropriateness and respect for the reader in its final documentation.

The evidence here is drawn from empirical and qualitative evaluations from two case studies and a single expert evaluator. Combining the indicators and comparing the data for these two projects suggested that social interactive adaptations and pro-social behaviour in a writing team's emails are reflected in the relational metadiscourse and social desirability of the document produced by the team. To prove this hypothesis scientifically and quantitatively requires repetition on multiple writing projects, preferably using a reader-focused evaluation method.

The methodology allows unintrusive research from both positivist and interpretive perspectives, to compare writing projects. With semi-automation of the process, the methodology can be used both to prove the concept that social interactive adaptations and pro-social behaviour in emails are reflected in the final document, and also to contribute more broadly to our understanding of team writing. Companies managing mediated communications for organizations have the technological infrastructures to support the software application of such a methodology, which may also be beneficial to discourse communities other than those of writing teams. With a sufficiently large data base of project communications and document evaluations, it may be possible to use simulation programs to study communication behaviour marker adaptations and document evaluations to identify an optimal social-task balance.

9. Conclusion

This research aimed to answer the question "Can we learn about the influence of team culture on virtual team writing from content analysis of email communications during projects?" I have explored the concept that adaptations in email metadiscourse reflect social-interactive adaptations in the final document. These adaptations provide a form of metric, a way of measuring the written communication behaviour of teams. Emails provide a proxy for studying communication behaviour, and together with outcome measures can contribute towards building a causal model of team writing to identify performance predictors. Emails carry knowledge on the task and social dimensions of team work and through analysis of their metadiscourse can contribute to an understanding of the social activity of the writing process. Not only the adaptations writers make for their readers, purpose and context of use, but also the levels of sociability, involvement and solidarity can be determined from email records. In depth analysis of adaptations together with participant feedback contributes to an understanding of team culture and dynamics and the underlying influences.

My first hypothesis explored whether it was possible to extract both writing influences, the independent variables, and communication behaviour markers, the dependent variables, from email content, and show dependencies, as predicted by the social interactive model of written communication:

> *Pilot study:*
>
> *H1 = Email communication behaviour is the product of writing influences and representative variables of both can be derived non-intrusively from email content.*

Being able to extract both types of data from emails provides an unintrusive methodology, without additional work for the researched, and no research or researcher influence in the context being researched. Such a methodology has potential for multiple studies of professional practice, without hindering work goals and deadlines. Multiple case studies, completed using a standard methodology, have the potential to provide generalized knowledge applicable to a wider field of professional practice. I analysed email data from a technical writing project in which I was the technical writer for this first part of my research. Analysis of communication markers representing effort and value of the communication, formality, involvement and solidarity all showed variations with writing influences in this study.

The second hypothesis explored whether interpersonal metadiscourse in emails might describe the social dimensions of networked team writing projects. The social contribution to team work is known to benefit performance, so that finding an unintrusive way to measure this aspect has the potential to predict performance. Such metadiscourse contributing to the social dimension requires research into form rather than content of emails and thus has potential for universal application across projects irrespective of the subject matter content.

> *H2 = Social dimensions of teams can be identified from email communications.*

I studied networked team writing projects from academic and commercial contexts for this part of my research and showed that empirical data extracted from email content could describe the social dimensions of team working. Email data interpretations were reinforced by participant perceptions, and in the commercial project reflected the power hierarchy of the business relationship; the contextual relationship between the writer and reader and organizational norms clearly influenced communication behaviour and email style. Social aspects identified in the academic project profiled administrative versus subject matter expertise functions of emails and the fact that there were two separate organizations involved, the Society and the Conference organizers. Email data profiled differences between the two contexts, such as higher pro-social communication behaviour in the academic context, use of pro-social communication strategies to build relationships in the academic project, and a more task-oriented focus in the commercial project. At a micro-level, in depth analyses with participant feedback helped to interpret changes in communication behaviour following changes in team composition, FtF meetings and milestones in projects.

Finally, with the same social interactive behaviour applied to written communication in emails and communication in the final document, and knowing the positive contribution of the social dimension of team work on performance, my third hypothesis explored whether social interactive adaptations in written emails with writing influences, and levels of pro-social communication behaviour might be reflected in the relational metadiscourse and social desirability of the final written document.

> *H3 = Social interactive adaptations and pro-social behaviour in a writing team's emails are reflected in the relational metadiscourse and social desirability of the document produced by the team.*

To address this final hypothesis, I compared social marker frequencies, correlations between pro-social communication strategies, task profiles and conclusions drawn in the case studies on the balance of social and task dimensions in the academic and commercial contexts. Social-task profiles were interpreted from social interactive adaptations of communication markers with writing influences. In the commercial project, involvement varied with purpose, but not with the reader, suggesting a more task oriented representation. In the academic project, writers adapted all the communication markers studied for their intended readers and of particular note, involvement varied with receiver and purpose, showing both task and social involvement. Comparing task profiles, however, did not suggest any less emphasis on task-oriented activities in the academic project. Mean frequencies for socio-emotional communication markers were higher in the academic project, and correlations between the markers suggested a higher socio-emotional component on the academic project. Of particular note in the commercial project was the absence of self-disclosure in communications. Feedback on solidarity also suggested a more highly profiled social dimension in the academic project. These comparisons showed a task orientation in the commercial context and an even social-task balance in the academic context. Document evaluations based on Sless's social desirability model showed a more appropriate relational content in the academic project, which paralleled the higher relational

content and social-interactive adaptations practised in the academic project. Thus in this study, social interactive adaptations and pro-social behaviour in emails during the projects were reflected in the final documents.

To prove the concept that social-interactive adaptations in metadiscourse and levels of pro-social communication behaviour are reflected in the final documents written by virtual teams requires repetition of this research in multiple projects. Concluding over his comparative discourse analysis of two genres in professional communication, CEO letters and directors' reports, Hyland suggests analysing metadiscourse in "high and low performing companies, good and bad years, or different commercial sectors". He suggests that such investigations:

> *...may reveal distinguishing features which operate in specific contexts. Further research in this area may reveal the ways writers control the expression of textual and interpersonal relationships within a text are as vital to the rhetorical success of a text as its propositional content (Hyland 1998 p241).*

In my research I have focused on relational markers in team communications, which have contributed to building relationships successfully and maintaining the team. In his conclusion Hyland argues that "Such studies can help learners gain a better understanding of the strategies used in corporate messages and develop a more effective rhetorical and verbal repertoire to use in the professional domains in which they will find themselves" (Hyland 1998 p242). Of particular interest to the concept tested in this research and to continue this line of research in real professional writing contexts, we need to research different corporate cultures which encourage or discourage personal email communications. My hypothesis is that the organizations whose cultures allow and encourage relational communications will show higher project performance.

The methodology has two professional applications. Firstly, detailed analysis of social-interactive adaptations by email authors to readers and reader-writer contexts provides a useful tool in monitoring the social dimensions of projects to identify and correct problem areas. Using member-checking where team members are shown analyses and significant variations in behaviours to develop interpretations has potential for identifying and correcting problem areas in long term writing projects. Multiple project comparisons need to be made to evaluate the consistency of behaviours of the markers against different influences, such as FtF meetings and team membership changes, to validate interpretations of social dynamics in teams and the underlying causes. In this way a database of trends can be collected for consistent interpretations of empirical data extracted from emails. With email monitoring, a kind of social-task email diagnostics could then be used to avoid problems in long term projects. Across context comparisons at this detailed level may also identify social norms linked to email as a genre rather than to a particular context of work or discourse community. For example, in this study, results from both the commercial and academic projects showed a decrease in involvement and social building units with increased audience size. Both projects also showed higher solidarity with an audience than without an audience, but lower solidarity in emails addressed to more than one individual.

Secondly, relating the social-task dimensions of projects to evaluations of performance is the first step towards identifying a causal model of virtual and semi-virtual team writing projects, to identify performance predictors. For example in the Keyware study, highest level managers showed the highest involvement. This observation might be worthy of study to assess what level of involvement at managerial level predicts good team performance. Falling below a certain involvement level might flag a potential problem.

Two further aims of this research were to develop an unintrusive methodology for researching writing projects and to research projects in a standard way to allow cross context comparisons and multiple case studies to deliver findings more broadly applicable to professional writing practice. This research has shown that systematic and consistent analysis of social-interactive writing behaviour on teams from communication markers in emails is possible, and does not require intrusion into the workplace. Analysis of emails provides a completely transparent way of researching writing projects. Additionally, with over 50% of communications in co-located teams being mediated (Mortensen and Hinds 2001), the methodology has potential for both distributed and co-located teams.

The methodology requires multiple communication markers, some of which are objective and some subjectively collected, thus providing exogenic and endogenic qualities and adopting a multiple realities interpretation of knowledge. Complexities in the representations of these markers require holistic interpretation to identify behaviour, and areas to improve our understanding of their representations emerged from this research, as I discuss here.

Formality was based on evaluations of isolated greeting and signature components. However, there was some evidence that interpretation of formality of one component might depend on the presence of another. Interpretation of formality may therefore be improved by a more contextualised holistic scoring system taking into account different combinations of greetings and salutations. An experimental design using emails with varying combinations of open, close, manual and automated signatures, and asking readers to assign single formality scores for the different combinations could be used to explore such a scoring system.

Politeness and general social building components of the social marker used in this research showed different trends with socialisation phase, suggesting varying underpinning strategies. The solidarity marker represented solidarity status, a strategy to build solidarity or representation of an official entity. Systematic protocols need to be developed to identify pro-social strategies and support interpretations of social building and solidarity markers as either social building or social maintenance strategies. Intentional strategies of different social building units such as apology, courtesy, self-disclosure etc. need to be studied and understood in depth to support consistent and reliable interpretations from their frequencies in emails.

In this research, I chose to focus on pro-social behaviour, because I feel there is a bias towards studying negative behaviour in email communications. Ducheneaut and colleagues distinguish 30 years of email research as having addressed three aspects: "email as a file cabinet extending human information processing capabilities, email as a production line and locus of

work coordination, and finally, email as a communication genre supporting social and organization processes" (Ducheneaut 2005 p11). They conclude, that the "research still fails to show where and how, in general, email will be used to good effect in organizations – if anything, it seems to imply that such guidance is impossible" (Ducheneaut 2005 p35). They identify the most problematic issue with email research as having been its failure to connect the three levels at which it operates: individual, communicative and the socio-organizational (Ducheneaut et al. 2005 p37). Ducheneaut and colleagues suggest that the behavioural components in email use connect the individual to each of these levels and that awareness of sender's normal behaviours would help receivers of emails.

> *Much more could be done to give email users a better sense of how to interact with their correspondents (Ducheneaut et al. 2005 p39).*
>
> *Overall it seems much attention has been focused on incidents and the problematic nature of email for decision making in organizations, rather than the countless rewarding and routine nonproblematic interactions also happening (Ducheneaut et al. 2005 p33).*

I believe that there is more to be learnt from the success of teams which demonstrate positive communication behaviours and have therefore not included social markers of negative behaviour in my research. I conclude on pro-social behaviour from adaptations and frequency of pro-social markers, rather than ratio of negative to positive social markers.

However, a project with high frequencies of social building units, according to Sproull and Kiesler's lack of social context cues theory (1986; see section 2.3.4), is also likely to have higher frequencies of other less inhibited behaviours, which are negative, such as flaming. Negative socio-emotional behaviour may not necessarily be detrimental to team performance, but markers of both positive and negative socio-emotional behaviour would provide a fuller picture of the professional team communication behaviours, which underpin successful projects. Future research therefore also needs to include markers of negative social behaviour to understand their influence on performance in networked team writing projects.

There was some evidence that FtF contact coloured an interviewee's perception of social behaviour in this research. The difference between perceptions of FtF social relations and perceptions of textually mediated social relations is an interesting area of study in itself. Another area for further study is the ratio of FtF meeting and CMC in projects and its impact on both the socio-emotional needs in CMC and project performance. This last is particularly important because real life writing teams involve combinations of FtF and virtual working. Variations in CMC behaviour with increasing FtF contact and the effects on performance of varying combinations of CMC and FtF communication therefore need to be researched. A final medium-related issue in this research was the potential influence of writer anticipation of FtF meeting with readers. The academic document studied in this research was designed for use in a FtF meeting between the author and reader. This writing influence, combined FtF and textual media delivery, may influence the social interactive adaptations contributing to a document's social appropriateness. This raises another interesting area for writing research; teams may

demonstrate more social interactivity in documents when they anticipate FtF meetings with readers.

Further in this research, the commercial project profiled English writers as more sociable, whereas the academic project profiled Dutch writers as more sociable. This concept needs exploring in multiple commercial and academic contexts to establish whether there is a language dependent difference in pro-social CMC behaviour between the two types of discourse community. Future research also needs to investigate the influence of gender balance in teams and influences on social interactive team behaviour.

The case studies in this research offer a snapshot of current day email communication behaviour in professional writing contexts. The merging of personal and professional styles which Danet (2001a) has predicted is already apparent in communications from the academic context in this research. Additionally, Danet and other researchers (see section 2.3.6) believe that email style is in evolution, so studies such as my own will have value for comparisons with future studies to describe the evolution. Exchange rate of emails also appears to influence email style, with faster exchanges adopting more conversational styles. Thus three time lines are relevant to email research, exchange rate, the socialisation phase with influences at different points in time, and the overall evolution of email.

The case studies reported here have shown that social and task elements are combined in teams, in individuals and within single exchanges without losing focus on the task. Professional email communications can and must offer more than a traditional business letter, because social dimensions and informal exchanges contribute positively to team performance. Emails provide the social construction of the networked team and must therefore provide both social and task dimensions and the opportunities for informal unplanned communication without inhibition or censure. To encourage informal exchange of ideas and improved socio-emotional relations, I recommend team members adopt a more conversational style in their exchanges. This will help to overcome social norms in commercial and other organizational contexts and to keep the door open for an informal exchange of ideas. I recognize my own formality in email communications and also my own reluctance to mix task and personal content. Through the kind permission of those who have participated in this research, I have gained an insight into a kind of professionalism which is not rule-bound by any kind of social norms, and which encompasses both dimensions, to the benefit of the team and the task.

This research has contributed towards validating the concept that content analysis of emails written on team projects can help us to understand the influences of team culture on virtual team writing. Adaptations of pro-social communication behaviours in emails to reader and writer-reader contexts have been used to identify the social-task profiles of teams and compare these profiles to evaluations of the final documents. The methodology developed offers a relatively unintrusive way to monitor ongoing projects, to identify and correct problem areas. Additionally, for writing research it offers a standard methodology for comparing multiple projects in a consistent way, so that results of writing research can be applied more generally.

Finally, the use of email communications in co-located teams reported in recent research suggests that the tool may also be applicable to co-located teams (Mortensen and Hinds 2001).

Glossary

I define here what I mean by the terms I use in this thesis.

Content analysis: In my research I have used what Busch et al. (n.d.) describe as conceptual content analysis. I have analysed texts to produce empirical data. Some interpretation of text units has been necessary and some numbers are derived from objective counts. Discourse has been analysed, although not in any of the discrete approaches which Phillips and Hardy (2002) list: interpretive structuralism, social linguistic analysis, critical discourse analysis or critical linguistic analysis.

Group maintenance issues: Any activity, which promotes group maintenance, including coordination, courtesy, pro-social behaviour, and which is not related to the group's primary task goal, e.g. writing a document.

Informality: I use informality and formality to describe the familiarity intended or interpreted from the style of the communication. Formal styles of communications infer formal and less familiar relationships. This is not to be confused with politeness. Formality and informality may both be used in polite or impolite communications, depending on the reader writer relationship and context of the communication. A formal email to a receiver in a relationship where the receiver is familiar to the sender may be interpreted as impolite, whereas a formal email to a receiver who is unfamiliar to the receiver may be appropriate and therefore interpreted as polite.

Locales: The locales theory of group behaviour (Greenberg et al. 1999) seeks to explain working groups and their use of artefacts, and describes "a social world as a group of people with some common purpose, a site for collaboration and some means to communicate" (p31). The locale may be physical or virtual, providing the site and means for the social world to collaborate. It is the "actual site in which a group collaborates, the actual means by which people communicate, and the actual means by which the work is achieved" (Greenberg et al. 1999 p32). A locale offers a foundation, a shared space with the tools and resources to support collaborative work and the awareness of others' work and changes in artefacts; interactions within the locales maintain a sense of shared place. Users need control of the locale, and the locales should help them co-ordinate and negotiate their work together (Greenberg et al. 1999 p 32; Noel and Robert 2003 p246). Users also need to view a locale or multiple locales from their own perspective according to their degree of involvement, manage and remain aware of evolving interactions over time and to relate different locales to one another (Greenberg et al. 1999 p32).

Networked team working: Any team activity which is at least in part by remote members who need to use mediated communications.

Relational messages / relational metadiscourse: As a metacommunicative function of interaction, relational messages are those verbal and nonverbal expressions that indicate how two or more people regard each other, regard their relationship, or regard themselves within the

context of the relationship (Burgoon and Hale 1984 p193). In my dissertation I use "relational messages" and "socio-emotional communication" to represent the same construct.

Socio-emotional communication: please see "relational messages" above.

Social dimension of team working: The social dimension of teams is basically concerned with the relationships, whether socio-emotional or organizational. The social dimension of team working is described by communications, activities and behaviours which are not addressed directly towards the group task goal, but rather towards relations, coordination of activities, welfare and maintenance of the team construct. In this research the team culture is described by the balance between social and task dimensions of team working.

Teams: For the purposes of this dissertation I define a team as any group of two or more professionals working towards a common goal. I do not differentiate between "collaborative" and "team" working.

Team culture: In this research, the team culture is described by the balance between social and task dimensions of team working.

Team writing: Team writing can involve two or more professionals working together to create a document. Team membership is not restricted to a particular contributing skill, such as writing, illustrating, administrating etc. If the member is contributing towards creation of the document, they are considered a collaborator or team member.

References

Adler, T.R., 2000: "An evaluation of the social perspective in the development of technical requirements", *IEEE Transactions on professional communication*, 43 (4) pp.386-396.

Anderson, P., 1985a: "What survey research tells us about writing at work", in L. Odell and D. Goswami, eds., *Writing in non-academic settings*, New York: Guilford, pp.3-83.

Anderson, P., 1985b: "Survey methodology", in L. Odell and D. Goswami, eds., *Writing in non-academic settings*, New York: Guilford, pp. 453-501.

Argyle, M., 1994: *The psychology of interpersonal behavior*, 5th ed., London: Penguin books.

Arts, A., Maes, A., Noordman, L. and Jansen, C., 2004: *Overspecification in instructive discourse*, [Presentation] Document Design Conference, Tilburg University, January 2004.

Bakardjieva, M. and Feenberg, A., 2001: "Involving the virtual subject," *Ethics and information technology*, 2 (4) pp.233-240.

Bales, R.F., 1951: *Interaction process analysis*, Cambridge (Mass.): Addison-Wesley Press Inc.

Bales, R. and Strodtbeck, F. 1951: "Phases in group problem-solving", *Journal of abnormal social psychology*, 46 (4), pp.485-495.

Barker, T. T., 1998: *Writing software documentation; a task-oriented approach,* Boston: Allyn & Bacon.

Barker, V. E., Abrams, J. R., Tiyaamornwong, V., Seibold, D. R., Duggan, A., Park, H. S. and Sebastian, M., 2000: "New contexts for relational communication in groups", *Small group research*, 31 (4), pp.470-503.

Baron, N., 1998: "Letters by phone or speech by other means; the linguistics of email", *Language and Communication*, 18 (2) pp. 133-170.

Baron, N., 2001: *Alphabet to Email: How Written English Evolved and Where It's Heading*: London: Routledge.

Barreto, M. and Ellemers, N., 2002: "The impact of anonymity and group identification on progroup behavior in computer-mediated groups", *Small group research*, 33 (5), pp.590-610.

Beaufort, A., 2000: "Learning the trade. A social apprenticeship model for gaining writing expertise", *Written communication*, 17 (2) pp.185-223.

Beck, E., 1993: "A survey of experiences of collaborative writing", in M. Sharples ed., *Computer supported collaborative writing*, London: Springer-Verlag, pp.87-112.

Beck, E. and Bellotti, V., 1993: "Informed opportunism as strategy: supporting coordination in distributed collaborative writing", in G. De Michelis, C. Simone and K. Schmidt, eds., *Proceedings of the Third European Conference on Computer-Supported Cooperative Work ECSCW '93, held in Milan, Italy, 1993*, pp.233-247.

Benne, K. and Sheats, P., 1948: "Functional roles of group members", *Journal of social issues*, 4 (2), pp.41-49.

Blaxter, L., Hughes, C. and Tight, M., 1996: *How to research*, Buckingham: Open University Press.

Bottitta, J., Prendergast, I. and Pappas, L., 2003: "Moving to single sourcing: managing the effects of organizational changes", *Technical communication*, 50 (3), pp.355-370.

Brown, P. and Levinson, S., 1987: *Politeness: some universals in language usage*, Cambridge: Cambridge University Press.

Burgoon, J. and Hale, J., 1984, "The fundamental topoi of relational communication", *Communication monographs*, 51 (3), pp.193-214.

Burgoon, J. and Hale, J., 1987: "Validation and measurement of the fundamental themes of relational communication", *Communication monographs*, 54 (1), pp.19-41.

Burke, K. and Aytes, K., 1998: "A longitudinal analysis of the effects of media richness on cohesion development and process satisfaction in computer-supported workgroups", in *Proceedings of the Thirty-first Hawaii International Conference on System Sciences HICSS*, IEEE Computer Press, 1, pp.134-144.

Burke, K. and Chidambaram, L. 1999: "How much bandwidth is enough? A longitudinal examination of media characteristics and group outcomes", *MIS Quarterly*, 23 (4), pp.557-580.

Burke, K., Chidambaram, L. and Aytes, K., 2002: "Do some things change faster than others? The dynamics of behavioral change in computer-supported groups," *Group decision and negotiation*, 11 (4), pp.293-309.

Burnett, R.E., 1997: *Technical Communication,* 4th ed., Belmont CA: Wadsworth Publishing Company.

Busch, C., De Maret, P., Flynn, T., Kellum, R., Le, S., Meyers, B., Saunders, M., White, R. and Palmquist, M., 2005: *Content Analysis*, [Online], Writing@CSU. Colorado State University Department of English. Available from http://writing.colostate.edu/guides/research/content/ [Accessed 24th August 2006].

Button, G. and Dourish, P., 1996: "Technomethodology; paradoxes and possibilities", *Proceedings of the Conference on Human Factors in Computing Systems, CHI 96, Vancouver*, BC Canada, April 13-18, 1996, pp.19-26. [Also online] Available from http://sigchi.org/chi96/proceedings/index.htm / under "Papers". [Accessed 24th August 2006].

Carliner, S., 1993: "A way with words – presenting information verbally", in C. M. Barnum and S. Carliner, eds., *Techniques for technical communicators,* Boston: Allyn and Bacon, pp138-186.

Carliner, S., 2000: "Physical, Cognitive, and Affective: A Three-Part Framework for Information Design", *Technical Communication*, 47 (4) pp.561-576.

Carliner, S., 2004a: *Keynote speech,* [Presentation] Document Design Conference, Tilburg University, January 2004.

Carliner, S., 2004b: "What do we manage? A survey of the management portfolios of large technical communication groups", *Technical Communication*, 51 (1), pp.45-67.

Carliner, S., 2006: "Current challenges of research in information design and document design," *in* S. Carliner, J. Verckens and C. de Waele, eds., *Recent research in information and document design*, Amsterdam: John Benjamins pp.1-24.

Chidambaram, L., 1996: "Relational development in computer supported groups", *Management Information Systems Quarterly*, 20 (2), pp.142-165.

Cole, P. and Nast-Cole, J., 1992: "A primer on group dynamics for groupware developers", in *Groupware: Software for computer-supported cooperative work*, D. Marca and G. Bock, eds.,

Los Alamitos CA: IEEE Computer Society Press.

Communication Research Institute of Australia (CRIA), 2005: [On line] Available from http://www.communication.org.au/html/about_us.html [Accessed August 2006].

Crystal, D., 2001: *Language and the Internet*, Cambridge: Cambridge University Press.

Daft, R. and Lengel, R., 1984: "Information richness: a new approach to managerial behaviour and organization design", *Research in organizational behaviour*, 6, pp.191-233.

Damrau, J., 2006: "Managing a virtual team", *Intercom Magazine of the Society for Technical Communication*, June 2006, pp.12-13.

Danet, B., 2001a: *Cyberpl@y: communicating online*, Oxford: Berg Publishers.

Danet, B., 2001b: *Studies of Cyberpl@y: Ethical and Methodological Aspects*, case study prepared for the Ethics Work Group, Association of Internet Researchers, September 2001. [Online] Available from http://pluto.mscc.huji.ac.il/~msdanet/papers/ethics2.pdf [Accessed 24th August 2006]. (The Final Report of the Work Group is referenced above under Ess, C. and AoIR and was approved Autumn 2002.)

David Sless, n.d.: [On line] Available from http://www.xplorcanada.org/media/central/SlessBio-2005.pdf#search=%22David%20Sless%22 [Accessed August 2006].

de Jong, M. and Schellens, P. J., 1997: "Reader-focused text evaluation. An overview of goals and methods", *Journal of Business and Technical Communication*, 11 (4), pp.402-432.

De Laet, J. and Obbels, W. 2006: *Introduction to the VIC HPC Cluster*, LUDIT, K.U.Leuven.

Dennis, A. R. and Kinney, S. 1998: "Testing media richness theory in the new media; the effects of cues, feedback and task equivocality", *Information Systems Research*, 9 (3), pp.256-274.

Denscombe, M., 1998: *The good research guide for small-scale social research projects*, Buckingham, Philadelphia: Open University Press.

Diaper, D., 1993: "Small-scale collaborative writing using electronic mail", in *Computer*

supported collaborative writing in practice, D. Diaper and C. Sanger, eds., London: Springer-Verlag, pp.69-92.

Dillon, A., 1993: "How collaborative is collaborative writing? An analysis of the production of two technical reports", in M. Sharples, ed., *Computer supported collaborative writing*, London: Springer-Verlag, pp.69-86.

Dillon, J. T., 1990: "The practice of questioning", in O. Hargie, ed., *International series on communication skills,* London and New York: Routledge.

Dorn, E., 1999: "Case method instruction in the business writing classroom", *Business communication quarterly*, 62 (1), pp.41-63.

Ducheneaut, N. and Watts, D., 2005: "In search of coherence: a review of email research", *Human-computer interaction*, 20 (1/2) pp.11-48.

Edwards, K., 2001: *Is collaborative work helping the writing process or hindering it? An investigation into the pros and cons of collaborative technical writing in a professional environment*, Master of Arts Thesis, School of Cultural Studies, Sheffield Hallam University.

Edwards, K., 2005a: "Research help needed", *ISTC Newsletter April 2005*, [Online] Available from http://www.istc.org.uk/Publications/Newsletter/News_2005/istcApr2005.pdf#page=2 [Accessed 24th August 2006].

Edwards, K., 2005b: "What kind of team work helps usability?" *The UPA Voice June 2005*, [Online] Available from www.upassoc.org/upa_publications/upa_voice/volumes/2005/june/teamwork_usability.html [Accessed 24th August 2006].

Edwards, K., Williams, N., Dujardin, A-F., Hartley, P. and Spaepen, A., 2004: *Developing an e-mail analysis tool for writing research*, [Presentation] Document Design Conference, Tilburg University, January 2004.

Edwards, K., Williams, N., Dujardin, A-F. and Spaepen, A., 2005: "What can email discourse and dynamics tell us about content creation?" in P. Masip and J. Rom, eds., *Proceedings of the III International Conference Communication and Reality. Digital Utopia in the Media: from discourses to facts. A Balance*, University Ramon Llull, Barcelona, May 2005.

Edwards, K., Williams, N., Dujardin, A-F. and Spaepen, A., 2006: "Developing an e-mail analysis tool for writing research", in S. Carliner, J. Verckens and C. de Waele, eds., *Recent research in information and document design*, Amsterdam: John Benjamins, pp.169-190.

Eggins, S. and Martin, J., 1997: "Genres and registers of discourse", in Van Dijk, T.A. ed., *Discourse studies: a multidisciplinary introduction, Vol. 1 Discourse as structure and process,* London: Sage Publications Inc. pp.230-256.

Ess, C. and the Association of Internet Researchers (AoIR), 2002: *Ethical decision-making and Internet research:* Recommendations from the AoIR ethics working committee, Approved by AoIR, November 27, 2002. [On line} Available from www.aoir.org/reports/ethics.pdf [Accessed 23rd August 2006].

European Commission, 2002: "eEurope benchmarking report says EU Internet penetration reaching plateau", *Cordis website*, [On line] Available from http://ica.cordis.lu/search/index.cfm?fuseaction=news.simpledocument&N_RCN=17965&CFID=106954&CFTOKEN=30628674 [Accessed 24th August 2006].

European Science Foundation, 2000:"Good scientific practice in research and scholarship", *European Science Foundation website*, [On line] Available from http://www.esf.org/sciencepolicy/170/ESPB10.pdf [Accessed 24th August 2006].

Evans, C. and Dion, K., 1991: "Group cohesion and performance. A meta-analysis.", *Small group research*, 22 (2) pp. 175-185.

Eveland, J. and Bikson, T. 1987: "Evolving electronic communication networks: An empirical assessment", *Office: Technology and people*, 3, pp.103-128.

Faigley, L., 1985: "Non-academic writing. The social perspective", in L. Odell and D. Goswami, eds., *Writing in non-academic settings*, New York: Guilford Press, pp.231-247.

Ferrara, K., Brunner, H. and Whittemore, G., 1991: "Interactive written discourse as an emergent register", *Written communication*, 8 (1), pp.8-34.

Fitzgerald, J., 1992: *Towards knowledge in writing; illustrations from revisions studies*, Berlin: Springer.

Flower, L. and Hayes, J., 1980: "The dynamics of composing: making plans and juggling

constraints", in L. W. Gregg and E. R. Steinberg, *Cognitive processes in writing*, eds., Hillsdale (N.J.): Lawrence Erlbaum Associates, pp.31-50.

Flower, L. and Hayes, J., 1981a: "Plans that guide the composing process", in C. H. Fredericksen and J. F. Dominic eds., *Writing; the nature development and teaching of written communication*, Hillsdale (N.J.): Lawrence Erlbaum, pp.39-58.

Flower, L. and Hayes, J., 1981b: "A cognitive process theory of writing", *College composition and communication*, 32 (4), pp.365-387.

Flyvbjerg, B., 2006: "Five misunderstandings about case-study research", *Qualitative inquiry*, 12 (2), pp.219-245.

Ford, K., Oberski, I. and Higgins, S., 2000: "Computer-aided qualitative analysis of interview data; some recommendations for collaborative working", *The qualitative report*, 4 (3&4) [On line] Available from http://www.nova.edu/ssss/QR/QR4-3/index.html [Accessed 24th August 2006].

Foth, M., 2006: "Network action research", *Action research*, 4 (2), pp.205-226.

Furnham and Jaspers, J., 1983: "The evidence for interactionism in psychology: a critical analysis of the situation-response inventories", *Personality and individual differences*, 4, (6), pp.627-644.

Gains, J., 1999: "Electronic mail - a new style of communication or just a new medium?: An investigation into the text features of e-mail", *English for specific purposes*, 18, (1), pp.81-101.

Gallivan, M. and Benbunan-Fich, R., 2005: "A framework for analyzing levels of analysis issues in studies of e-collaboration", *IEEE Transactions on professional communication*, 48, (1), pp.87-103.

Geisler, C. Bazerman, C., Doheny-Farina, S., Gurak, L., Haas, C., Johndan, J-E., Kaufer, D.S. Lunsford, A., and Miller, C.R.,: 2001: "IText: Future directions for research on the relationship between information technology and writing", *Journal of business and technical communication*, 15 (3), pp.269-308.

Gersick, C. 1988: "Time and transition in work teams: toward a new model of group development", *Academy of Management Journal*, 31 (1), pp.9-41.

Gersick, C., 1989: "Marking time; predictable transitions in task groups", *Academy of Management Journal*, 32 (2), pp.274-309.

Goldstein, J. and Sabin, R. 2006: "Using speech acts to categorize email and identify email genres", in *Proceedings 39th Hawaii International Conference on system sciences*, 3, p.50. [On line] Available from http://csdl2.computer.org/comp/proceedings/hicss/2006/2507/03/250730050b.pdf [Accessed 24th August 2004].

Greenberg, S., Fitzpatrick, G., Gutwin, C. and Kaplan, S. 1999: "Adapting the locales framework for heuristic evaluation of groupware", *Proceedings of the Annual Conference of the Australian Computer-Human Interaction Special Interest Group, OzCHI '99*, pp.30-36.

Greene, J. and D'Oliveira, M., 2006: *Learning to use statistical tests in psychology*, 3rd ed., Maidenhead: Open University Press.

Gutwin, C. and Greenberg, S., 2002: "A descriptive framework of workspace awareness for real-time groupware," *Computer supported cooperative work*, 11, Netherlands: Kluwer Academic Publishers pp.411-446.

Hak, T., Willimack, D., and Anderson, A. E., 2004: *Ethnographic methods for evaluating document design; the use of onsite observation and onsite interviews in testing establishment survey questionnaires*, [Presentation] Document Design Conference, Tilburg University, January 2004.

Harrison, S., 2000a: *The discourse structure of email communications*, PhD thesis, University of Central England.

Harrison, S. 2000b: "Maintaining the virtual community: use of politeness strategies in an email discussion group", in L. Pemberton and S. Shurville, eds., *Words on the Web*, pp.69-78.

Harrison, S., 2004: "Reasoned argument or gossip? Metadiscussion in an email discussion group", in L. Allal and J. Dolz, eds., *Proceedings of the 9th International Conference of the EARLI Special Interest Group on Writing*, University of Geneva, September 2004, [CD].

Hart, G. 2004: "Practical and effective metrics", *Intercom Magazine of the Society for Technical Communication*, February 2004, pp.6-8.

Hartley, J., 1994: *Designing instructional text*, 3rd ed., Kogan Page.

Hartley, J., 2004: *Applying psychology to text design: A personal view*, [Presentation] Document Design Conference, Tilburg University, January 2004.

Hartley, P., 1997: *Group communication*, London and New York: Routledge.

Haugen, D., n.d.: *Editors, Rules, and Revision* [Online] Available from http://www.wcdd.com/dd/articles/STCintro.html [Accessed 12th September 2006].

Hayes, J., 1996: "A new framework for understanding cognition and affect in writing", in C. M. Levy and S. Ransdell, eds., *The science of writing. Theories, methods, individual differences, and applications*, 1996, Mahwah (N.J.): Lawrence Erlbaum Associates, pp.1-27.

Hayes, J. and Flower, L., 1980: "Identifying the organization of writing processes", in L. W. Gregg and E. R. Steinberg, eds., *Cognitive processes in writing*, Hillsdale (N.J.): Lawrence Erlbaum, 1980, pp.3-30.

Heylighen, F., 1998: "Basic concepts of systems approach", *Principia Cybernetica website*, [Online], Available from: http://pespmc1.vub.ac.be/SYSAPPR.html [Accessed 24th August 2006].

Honeycutt, L., 2001: "Comparing e-mail and synchronous conferencing in online peer response", Written Communication, 18 (1), pp.26-60.

Horton, W., 1993: "Pictures please – Presenting information visually", in C. M. Barnum and S. Carliner, eds., *Techniques for technical communicators,* Boston: Allyn and Bacon, pp187-218.

Hyland, K., 1998: "Exploring corporate rhetoric: metadiscourse in the CEO's letter", *The Journal of Business Communication*, 35 (2), pp.224-245.

Janssen, D., van Waes, L. and van den Bergh, H., 1996: "Effects of thinking aloud on writing processes", in C. M. Levy and S. Ransdell, eds., *The science of writing. Theories, methods, individual differences, and applications*, 1996, Mahwah (N.J.): Lawrence Erlbaum Associates, pp.233-250.

Kelly, L. and Duran, R., 1985: "Interaction and performance in small groups: a descriptive report", *International Journal of Small Group Research*, 1, pp.182-192.

Kendrick, J., 1998: "Group interviews as source for writing proposals," [On line] Available from http://users.erols.com/p2c2/nov98nws.html [Accessed 25th February 2001].

Kim, H. E. and Eklundh, K. S., 2001: "Reviewing practices in collaborative writing", *Computer supported cooperative work*, 10, pp.247-259.

Kim, Y., Hiltz, S. R. and Turoff, M., 2002: "Coordination structures and system restrictiveness in distributed group support systems", *Group decision and negotiation*, 11 (5), pp.379-404.

Kirkman, J., 1992: *Good Style: writing for science and technology*, London: E & FN Spon.

Kraut, R., Egido, C. and Galegher, J., 1988: "Patterns of contact and communication in scientific research collaboration", in *Proceedings of the 1988 ACM conference on Computer-supported cooperative work*, New York: ACM Press, pp.1-12.

Larbi, N. and Springfield, S., 2004: "When no one's home: being a writer on remote project teams", *Technical Communication*, vol. 51(1), pp. 102-108.

Lazega, E. and Pattison, P. E., 1999: "Multiplexity, generalized exchange and cooperation in organizations; a case study", *Social Networks*, 21, pp.67-90.

Lea, M. and Spears, R., 1992: "Paralanguage and social perception in computer-mediated communication", *Journal of organizational computing*, 2 (3&4), pp.321-341.

Lentz, L. and de Jong, M., 1997: "The evaluation of text quality: expert-focused and reader-focused methods compared", *IEEE Transactions on professional communication*, 40 (3), pp.224-234.

Lunsford, A. and Ede, L., 1990: *Singular texts/plural authors; perspectives on collaborative writing*, Southern Illinois University Press.

MacNealy, M.S., 1999: *Strategies for empirical research in writing*, Massachussetts: Allyn & Bacon.

Marca, D. and Bock, G., eds., 1992: *Groupware: Software for computer-supported cooperative work,* Los Alamitos CA: IEEE Computer Society Press.

McGee, L., 2000: "Communication channels used by technical writers throughout the

documentation process", *Technical Communication*, 47 (1), pp.35-50.

Miller, D.C., 1991: *Handbook of research design and social measurement*, 5th ed., Newbury Park (C.A.): Sage Publications.

Mitchell, A., 1996: *Communication and shared understanding in collaborative writing*, Master of Science Thesis, Toronto University, [On line] Available from http://home.pacific.net.sg/~kamitchell/thesis/thesis.pdf [Accessed 24th August 2006].

Mitchell, A., Posner, I. and Baecker, R., 1995: "Learning to write together using groupware", in I. R. Katz, R. Mack, L. Marks, M. B. Rosson, and J.Nielsen, eds., *CHI 95: Conference on Human Factors in Computing Systems*, New York: Association for Computing Machinery, Inc.

Mortensen, M. and Hinds, P., 2001: "Conflict and shared identity in geographically distributed teams", *The international journal of conflict management*, 12, (3), pp.212-238.

Mulholland, J., 1999: "E-mail: uses, issues and problems in an institutional setting", in F. Bargiela-Chiappini and C. Nickerson, eds., *Writing business: genres, media and discourses*, pp.57-84.

Mulvey, P. W. and Klein, H. J., 1998: "The impact of perceived loafing and collective efficacy on group goal processes and group performance," *Organizational behavior and human decision processes*, 74, 62-87.

Narayanan, A. and Watts, D., n.d.: *Exact Methods in the NPAR1WAY Procedure*, [Online], Available from: http://www.ats.ucla.edu/STAT/SAS/library/exact.pdf [Accessed 12/12/2005].

Newman, J. and Newman, R., 1992: "Three modes of collaborative writing", in P. O'Brian Holt and N. Williams, eds., *Computers and writing: state of the art,* pp.20-28.

Nickerson, C., 2000: *Playing the corporate language game*, Amsterdam: Rodopi BV.

Noel, S. and Robert, J., 2003: "How the Web is used to support collaborative writing", *Behavior and Information Technology*, 22 (4), pp.245-262.

Nord, M. A. and Tanner, B., 1993: "Design that delivers – formatting information for print and online documents" in C. M. Barnum and S. Carliner, eds., *Techniques for technical communicators,* Boston: Allyn and Bacon, pp.219-251.

Nystrand, M., 1982a: "An analysis of errors in written communication" in Nystrand, M., ed., *What writers know: the language, process, and structure of written discourse*, New York: Academic Press, pp.57-74.

Nystrand, M., 1982b: "The structure of textual space," in Nystrand, M., ed., *What writers know: the language, process, and structure of written discourse*, New York: Academic Press, pp.75-86.

Nystrand, M., 1986: *The structure of written communication. Studies in reciprocity between writers and readers*, Orlando, Florida: Academic Press.

Nystrand, M., 1989: "A social-interactive model of writing", *Written Communication*, 6, (1), pp.66-85.

Nystrand, M., 1990: "Sharing words", *Written Communication*, 7, (1), pp3-24.

Odell, L., 1985: "Beyond the text; relations between writing and social context," in L. Odell and D. Goswami, eds., *Writing in non-academic settings,* New York: Guilford Press, pp249-280.

Panteli, N., 2004: "Discursive articulations of presence in virtual organizing", *Information and Organization*, 14 (1), pp. 59-81.

Paradis, J., Dobrin, D. and Miller, R., 1985: "Writing at Exxon ITD: notes on the writing environment of an R&D organization", in L. Odell and D. Goswami, eds., *Writing in non-academic settings*, New York: Guilford, pp.281-307.

Pauleen, D. and Yoong, P., 2001: "Facilitating virtual team relationships via Internet and conventional communication channels", *Internet research: electronic networking applications and policy*, 11 (3), pp.190-222.

Pendharkar, P. and Young, K., 2004: "The development of a construct for measuring an individual's perceptions of email as a medium for electronic communication in organizations", *IEEE Transactions on professional communication*, 47 (2), pp.130-143.

Perrin, D., 2003: "Progression analysis PA: investigating writing strategies at the workplace", *Journal of Pragmatics*, 35, pp.907-921.

Perrin, D. 2005: *Tracing journalists' language awareness*, [Presentation] Language in the media: representations, identities, ideologies, Leeds University, School of modern languages

and cultures: linguistics & phonetics, September 2005.

Petelin, R. 2002: "Managing organisational writing to enhance corporate credibility", *Journal of Communication Management*, 7 (2), pp.172-180.

Phillips, N. and Hardy, C., 2002: *Discourse analysis: investigating processes of social construction*, Thousand Oaks, California: Sage Publications.

Plain English Campaign, 1993: *The plain English story*, 3rd ed., Stockport: Plain English Campaign.

Plowman, L., Sharples, M. and Goodlet, J., 1993: *The development of a cognitive model for computer support of collaborative writing: end of project report*, University of Sussex Cognitive Science Research Papers, 9.

Porter, L. R. and Coggin, W., 1995: *Research strategies in technical communication*, New York: Wiley.

Postmes, T., Spears, R. and Lea, M., 1998: "Breaching or building social boundaries? SIDE-effects of computer-mediated communication", *Communication Research*, 25 (6) pp.689-715.

Postmes, T., Spears, R. and Lea, M., 2000: "The formation of group norms in computer-mediated communication", *Human Communication Research*, 26 (3), pp.341-371.

Preston, A., 2004: "Types of usability methods", *Usability Interface Society of Technical Communication*, January 2004, p.15-15.

Roberts, C., 2001: "Content analysis", in N. J. Smelser and P. Baltes, eds., *International encyclopaedia of the social and behavioural sciences*, Elsevier, pp.2697-2702.

Rooksby, E., 2002: *E-mail and ethics: Style and ethical relations in computer-mediated communications*, London: Routledge.

Root, R., 1988: "Design of a multi-media vehicle for social browsing", in *Proceedings of the conference on computer supported cooperative work CSCW '88*, pp.25-38.

Rude, C., 2002: *Technical editing*, 3rd ed., Allyn & Bacon.

SAS Institute Inc., n.d.: *NPAR1WAY Procedure, Exact Tests, Computational Resources*,

[Online] Available from http://technology.msb.edu/training/Statistics/SAS/books/stat/chap42/sect19.htm [Accessed 22nd August 2006].

SAS Institute Inc., n.d.: *Sample 507: Compute estimates and tests of agreement among multiple raters,* [Online] Available from http://support.sas.com/ctx/samples/index.jsp?sid=507&tab=details [Accessed 22nd August 2006].

SAS Institute Inc., n.d.: *Sample 528: %VARMOD macro to fit a normal model accommodating unequal variances*, [Online] Available from http://support.sas.com/ctx/samples/index.jsp?sid=528 [Accessed 24th August 2006].

SAS Institute Inc., n.d.: *SAS/STAT Users Guide*, [Online] Available from http://v8doc.sas.com/sashtml/ [Accessed 22 August 2006].

Schmitz, J. and Fulk, J., 1991: "Organizational colleagues, media richness, and electronic mail", *Communication Research*, 18 (4) pp.487-523.

Schriver, K. A., 1997: *Dynamics in document design*, New York: John Wiley & Sons, Inc.

Sharma, S., 1996: *Applied multivariate techniques*, New York: John Wiley & Sons, Inc.

Sharples, M. 1993: "Adding a little structure to collaborative writing", in D. Diaper and C. Sanger, eds., *Computer supported collaborative writing in practice*, London: Springer-Verlag, pp.51-68.

Sharples, M., 1996a: "Designs for New Writing Environments," in M. Sharples and T. van der Geest, eds., 1996: *The new writing environment: writers at work in a world of technology*, London: Springer-Verlag, pp.97-113.

Sharples, M., 1996b: "An account of writing as creative design", in C. M. Levy and S. Ransdell, eds., *The science of writing. Theories, methods, individual differences, and applications*, 1996, Mahwah (N.J.): Lawrence Erlbaum Associates, pp.127-148.

Sharples, M., 1999a: *CommonSpace as a general purpose writing environment*, [On line] Available from http://www.sixthfloor.com/sharples.html [Accessed 25th October 2000].

Sharples, M., 1999b: *How we write*, London: Routledge.

Sharples, M. and Pemberton, L., 1988: *Representing writing: an account of the writing process with regard to the writers' external representations,* University of Sussex Cognitive Science Research Papers, 119.

Sharples, M. and Pemberton, L., 1990: *Starting from the writer; guidelines for the design of user-centered document processors*, University of Sussex Cognitive Science Research Papers, 154.

Sharples, M., Goodlet, J. S. and Pemberton, L., 1989: "Developing a writer's assistant", in P. Holt and N. Williams, eds., *Computers and writing: models and tools*, Oxford: Blackwell, pp.22-37.

Sharples, M., Goodlet, J. S., Beck E. E., Wood, C. C., Easterbrook, S. M. and Plowman, L., 1992: *Research issues in the study of computer supported collaborative writing*, University of Sussex Cognitive Science Research Papers, 119, also in M. Sharples, ed., *Computer supported collaborative writing,* London: Springer-Verlag pp.9-28.

Sheffield Hallam University, 2005: *Research ethics policies and procedures*, 2nd ed., [On line] Available from http://www.shu.ac.uk/research/docs/ethics_policy_2004.pdf [Accessed 24th August 2006].

Sherblom, J., 1988: "Direction, function and signature in electronic mail", *Journal of business communication*, 25 (4), pp.39-54.

Slater, P., 1965: "Role differentiation in small groups", in A. P. Hare, E. Borgatta and R. Bales, eds., *Small Groups*, Revised ed., New York: Knopf pp.610-627.

Sless, D., 2004: "Designing public documents", *Information Design Journal + Document Design*, 12 (1), pp. 24-35.

Sproull, L. and Kiesler, S., 1986: "Reducing social context cues: electronic mail in organizational communication", *Management Science*, 32 (11), pp.1492-1512.

Strauss, A. L. and Corbin, J., 1991: *Basics of Qualitative Research*, Newbury Park, California: Sage Publications Inc.

Strauss, A.L., 1987: *Qualitative analysis for social scientists*, Cambridge: Cambridge University Press.

Suchman, L., 1987: "Situated actions", in Suchman, L., *Plans and situated actions; the problem of human machine communication*, Cambridge University Press, pp.49-67.

Taylor, J., 2000: "Electronic mail, communication and social identity: a social psychological analysis of computer-mediated group interaction", in L. Pemberton and S. Shurville, eds., *Words on the Web*, Exeter, England: Intellect™, pp.96-78.

Te'eni, D., Sagie, A., Schwartz, D. G., Zaidman, N. and Amichai-Hamburger, Y., 2001: "The process of organizational communication: a model and field study", *IEEE Transactions on Professional Communication*, 44 (1), pp.6-20.

Thurlow, C., Lengel, L. and Tomic, A., 2004: *Computer mediated communication: Social interaction and the Internet*, London: Sage Publications.

Tidwell, L. and Walther, J., 2002: "Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations. Getting to know one another a bit at a time", *Human Communication Research*, 28 (3), pp.317-348.

Timmerman, E. and Scott, C., 2006: "Virtually working: communicative and structural predictors of media use and key outcomes in virtual work teams", *Communication Monographs*, 73 (1), pp.108-136.

Tucker, R. and Panteli, N., 2003: "Back to basics: Sharing goals and developing trust in global virtual teams", in N. Korpela, R. Montealegre and A. Poulymenakou, eds., *Organizational information systems in the context of globalization*, Boston: Kluwer Academic Publishers, pp.85-98.

Tufte, E.R., 1997: *Visual Explanations: Images and quantities, evidence and narrative*, Graphics Press.

Turchin, V., 1991: "Reflection correspondence theory", *Principia Cybernetica website*, [Online], Available from: http://pespmc1.vub.ac.be/REFCORR.html [Accessed 24th August 2006].

Turk, C., and Kirkman, J., 1989: *Effective writing*, London: E. & F. Spon.

UCLA Academic Technology Services, n.d.: *What statistical analysis should I use? Statistical analyses using SAS,*[On line] Available from http://www.ats.ucla.edu/stat/sas/whatstat/whatstat.htm#kw [Accessed 22 August 2006].

Vaes, J., Paladinao, M. P. and Leyens, J. P., 2002: "The lost e-mail; pro social reactions induced by uniquely human emotions", *British Journal of Social Psychology*, 41 (4), pp.521-534.

Van Buren, R., and Buehler, M.F., 1980: *The Levels of Edit*, 2nd ed., JPL Publication 80-1 Pasadena, California: California Institute of Technology.

van der Geest, T., 1996a: "Chapter 1: Professional writing studied: authors' accounts of planning in the document production processes", in M. Sharples and T. van der Geest, eds., *The new writing environment: writers at work in a world of technology*, London: Springer-Verlag, pp.7-24.

van der Geest, T., 1996b: "Studying "Real Life writing processes: a proposal and an example", in C. M. Levy and S. Ransdell, eds., *The science of writing. Theories, methods, individual differences, and applications*, 1996, Mahwah (N.J.): Lawrence Erlbaum Associates, pp.309-322.

van der Meij, H., de Vries, B., Boersma, K., Pieters, J. and Wegerif, R., 2004: "An examination of interactional coherence in email use in elementary schools", *Computers in Human Behavior*, 21 (3), pp.417-439.

Walczuch, R. and Watson, R., 2001: "Analyzing group data in MIS research: including the effect of the group", *Group Decision and Negotiation*, 10 (1), pp.83-94.

Walther, J. B., 1992: "A longitudinal experiment on relational tone in computer-mediated and face to face interaction", in *Proceedings of the Twenty-fifth Hawaii international conference on system sciences*, 1992, IEEE, 4, pp.220-231.

Walther, J. B., 1995: "Relational aspects of computer-mediated communication: experimental observations over time", *Organization Science*, 6 (2), pp.186-203.

Walther, J. B. and Parks, M., 2002: "Cues filtered out: cues filtered in: Computer-mediated communication and relationships", in M. Knapp and J. Daly, eds., *Handbook of interpersonal communication*, Thousand Oaks, California: Sage, pp.529-563.

Williams, N., 1992: "Computers and Writing", in C. Butler, ed., *Computers and Texts: An Applied Perspective*, Oxford: Blackwell, pp.247-265.

Wilson, G.L., 1986: *Groups in context*, 4th ed., New York: McGraw-Hill, Inc.

Wood, C. C. 1992: *A cultural-cognitive approach to collaborative writing*, University of Sussex Cognitive Science Research Papers, 242.

Wuensch, K. 2003: *Inter-rater agreement*, [Online] Available from http://core.ecu.edu/psyc/wuenschk/docs30/InterRater.doc [22 August 2006].

Yates, J. and Orlikowski, W., 1993: *Knee-jerk anti-LOOPism and other email phenomena: oral, written and electronic patterns in computer-mediated communication.*, [Online] Available from http://ccs.mit.edu/papers/CCSWP150.html [Accessed June 2006].

Yates, S. 1996: "Oral and written linguistic aspects of computer conferencing: a corpus based study", in S. Herring, ed., *Computer-Mediated Communication. Linguistic, social and cross-cultural perspectives,* Amsterdam: John Benjamins, pp.29-46.

Yates, S.J. and Sumner T.R., 1997: "Digital Genres and the New Burden of Fixity", in Sprague, R.H., ed., *Proceedings of the Thirtieth Hawaiian International Conference on Systems Sciences, Vol. VI,* IEEE Computer Society Press pp.3-12.

Appendix A: List of variables, which influence the writing process (Edwards 2001)

Textual perspective:

Purpose of document (Dillon 1993)

Timescale for production and/or availability of resource to create document (Dillon 1993)

Influence of other work on this work (Dillon 1993)

Subjects to include or avoid (Dillon 1993)

Content organization (Faigley 1985)

Readability (Faigley 1985)

Appropriateness for audience (Faigley 1985)

Topic progression and flow (Faigley 1985)

Style (Faigley 1985)

Individual perspective

Individual's cognitive processes (Flower and Hayes 1980; Faigley 1985)

Strategies writers use (Faigley 1985)

How writers make certain choices (Faigley 1985; Odell 1985)

Writer's goals in composing (Nystrand 1989)

How writers' formulation directs production of resulting text (Faigley 1985)

Subject knowledge (Flower and Hayes 1980)

Familiarity with audience (Flower and Hayes 1980)

Social aspects

Context for production (Faigley 1985; Mitchell 1996; Nystrand 1989; Odell 1985)

Context for reception (audience) (Nystrand 1989)

Previous texts (Faigley 1985; Odell 1985)

Audience purposes (Faigley 1985; Nystrand 1989)

Negotiation between reader and writer (Nystrand 1989)

Reader interpretation (Nystrand 1989)

Hierarchy of text value relative to goals (Beaufort 2000)

Collaboration

Discourse, organizational, functional, or other group norms and hierarchies of norms (Adler 2000; Odell 1985)

How individual acts of communication define, organize and maintain social groups (Faigley 1985)

Social roles (Faigley 1985; Sharples 1999b)

Group purposes (Faigley 1985)

Communal organization (Faigley 1985; Odell 1985)

Ideology and culture (Faigley 1985; Odell 1985)

Social interactions (Faigley 1985; Beaufort 2000)

Internalised values, attitudes, knowledge ways of acting (Odell 1985)

Continued over ⇨

Collaboration

Group goals (Faigley 1985)

Whether people have worked together before (established norms) (Ford et al. 2000)

Need for consensus and/or negotiation (Beck 1993; Mitchell 1996; Dillon 1993; Diaper 1993)

Interrelationships: e.g. changes in style as influenced by each other (Beck 1993)

Process of text production (Dillon 1993)

Whether worked collaboratively at all before (Diaper 1993; Ford 2000)

Familiarity of context (socialisation) (Beaufort 2000)

Remote or local geographical location (Diaper 1993)

Individual's knowledge of subject matter, writing experience, context specific knowledge, interpersonal and analytical skills (Odell 1985; Beaufort 2000;)

Group composition (Beck 1993)

Political structure of team: power, status and hierarchy (Adler 2000; Dillon 1993)

Time availability (Beck 1993)

Ownership issues (Sharples 1993; Dillon 1993)

Asynchronous or synchronous collaboration (Sharples 1993; Mitchell 1996)

Supporting technology for document creation and communication (Sharples 1993; Diaper 1993)

Leadership assignment (Dillon 1993; Beck 1993)

Responsibility for production (Dillon 1993)

Social roles Human nature and interrelationships (Mitchell 1996; Sharples 1999; Beck 1993; Kendrick 1998; Diaper 1993)

Communal organization (Faigley 1985; Odell 1985)

Ideology and culture (Faigley 1985; Odell 1985)

Differences in age, expertise, experience, aspiration, gender, research backgrounds, education and interests (Ford et al. 2000)

Differences in technology skills (Ford et al. 2000)

Changing group membership (Beck 1993)

Work allocation and coordination (Sharples 1993; Beck 1993)

Role adoptions and changes (Sharples 1993)

Interdependence (Sharples 1993)

Management of conflict (Sharples 1993)

Communication issues (Sharples 1993; Diaper 1993)

Appendix B: Group Variables

Group Variables	Explanation	Notes
Purpose or goal (Faigley 1985)	The common purpose that defines the group.	Level of challenge (Hartley 1997): level of difficulty of the goal relative to group ability. Difficult goals challenge and result in higher performance (Hartley 1997 p142). Specificity of goal: how well the goal is defined for the group. Research has shown that group effectiveness is affected by clarity and definition of group goals (Hartley 1997 p142), across a wide variety of situations and goals. Explicit goals help with collaborative writing (Edes and Lunsford cited in Mitchell 1996).
Group size	The number of members in the group.	Ringelmann effect (Hartley 1997 p135). There is an inverse relationship between the number of people and individual performance in group tasks. The size of the group imposes a restriction on the volume of work which can be achieved by the specified deadline. Work-groups should not be larger than 15 (Argyle 1994 p175).
Group age (Bottitta et al. 2003)	How long the group has existed.	The age of the group has a direct effect on the group's stability and behaviour (Bottitta et al. 2003).
Group stability (Beck 1993; Bottitta et al. 2003)	Whether the group is permanent or temporary and expectations over its longevity. Whether group membership is continually changing or relatively stable.	Effects on establishment of norms, group evolution through phases, socialisation of new members. Hemphill's index of group dimensions (Miller 1991 p366) looks at certain characteristics of groups, such as stability.
Membership composition (Bakardjieva and Feenberg 2000; Ford et al. 2000)	The combined characteristic variables of the individual group members (see below under individual variables (Appendix C).	Given the individual characteristic values, a combined score could be determined as a group score indicative of group balance. "In a group, balance and flexibility are the keys" (Cole and Nast-Cole 1992 p49).

Continued over

Group Variables	Explanation	Notes
Role composition (optimally equilibrium) (Beck 1993)	A role is the stance a group member adopts in a group, for any period of time and is not related to job titles. Roles are changeable and related to social context. The role composition describes the communal structure (Faigley 1985; Odell 1985) of the group, in the sense of leadership and other roles of the team members etc.	Leadership (optional) (Beck 1993; Dillon 1993). A group leader may be formally designated, elected or evolve naturally (Beck 1993). Kantor defines four roles, mover, opposer, follower and bystander (Kantor 1975 cited in Cole and Nast-Cole 1992 p49). "In a group, balance and flexibility are the keys" (Cole and Nast-Cole 1992 p49).

Continued over

Group Variables	Explanation	Notes
Level of cohesiveness (Hartley 1997; Wilson 1986)	The level of cohesiveness is a measure of the closeness of members of the group, or team spirit; how well the members work and interact together.	Productivity increases with group cohesiveness peaks and then declines, i.e. there is an optimum and the relationship between productivity and group cohesiveness is not linear. (Argyle 1994; Wilson 1986). Too much social activity at the expense of work has a detrimental effect. Another disadvantage of high cohesiveness can be hostility towards other groups (Argyle 1994). Groups become cohesive through: ☐ physical proximity e.g. sharing an office; ☐ similarity of background, values, interests; ☐ rewarding experiences in the group, including success at tasks; ☐ activities which require co-operation; ☐ a skilled leader, who can preserve harmony; ☐ the absence of hostile or disturbed personalities (Argyle 1994). Cohesiveness is important because members cooperate and help each other better, enjoy the group, have higher job satisfaction, lower absenteeism and turnover and greater productivity at tasks (Argyle 1994 p167-8). Low cohesiveness leads to little co-operation, output falls and absenteeism and labour turnover increase. Levels of cohesiveness of a group have been interpreted using Bales second Interactive Process Analysis, called SYMLOG. This technique characterizes activities of groups on three dimensions: productivity towards purpose, friendly or unfriendly and dominant or submissive behaviours (Bales 1980, cited in Hartley 1997 p37). Shambaugh describes group development as driven by two dimensions, psychological closeness of the group and the accumulation of group culture (Shambaugh 1978 cited in Hartley 1997 p63). Tajfel and Fraser's group dimensions to measure "groupness" are interaction, perception, goals, norms, roles and affectation (Tajfel and Fraser 1978 cited in Hartley 1997 p23). Group cohesiveness has also been measured through perceptions (Aytes et al. 2002). Seashore's group cohesiveness index is based on responses to 3 questions (Miller 1991 p375). Level of follower acceptance to leader might also be an indicator of group cohesiveness (Hartley 1997 p 93). Group cohesiveness should be high (Argyle 1994 p175).

Continued over

Group Variables	Explanation	Notes
Subgroup existence (formal, informal, sociometric) (Hartley 1997; Lazega and Pattison 1999)	Groupings of 2 or more people identifiable within a group, based on professional interactions or "liking".	Subgroup existence may be motivated by liking/affiliation, or based on professional interactions. Liking has been measured by Moreno (Miller 1991; also cited in Hartley 1997 p128). His sociometry scale of spontaneous choice and sociometric preference is reported in Miller (Miller 1991 p376). The scale identifies friendship networks and is based on judgements on liking of other members and ranking according to preferences to work with them. Lageza et al. (Lazega and Pattison 1999) measured networking tendencies in formal, informal and friendly interactions. Formal and informal professional associations were defined as within or outside the designated working group. Informal associations were when partners working on a legal case asked colleagues other than those assigned to the case for informal advice. Affective relations are interpreted as one of the six group characteristics by Tajfel and Fraser: interaction, perception, goals, norms, roles, affective relations (Tajfel and Fraser 1978 cited in Hartley 1997 p23).
Norms (functional, discourse, organizational/hierarchical) (Adler 2000; Odell 1985)	Group culture and behaviour: Behaviours adopted and accepted as normal by a group which may be defined by discourse community, function, organization, or "friendship".	All groups form "norms" (Argyle 1994 p166). Organizational or hierarchical norms affect the interpersonal behaviour between, for example, a secretary and Managing Director. Group pressure (Hartley 1997 p118). "The risky shift" - tendencies for group decisions to be more extreme than those of the individuals (Hartley 1997 p148; Argyle 1994 p 173). Influence of hierarchy (organizational norms) on direction of communication. Two thirds of communication is in an upwards direction hierarchically (Wilke 1994 cited in Hartley 1997 p127). Ringelmann effect (Hartley 1997 p135). There is an inverse relationship between the number of people and individual performance in group tasks. Social loafing in groups: "the more people, the less effort they put in" (Argyle 1994 p172; Mulvey and Klein 1998). Group performance out-performs the average member of the group, but not the best member (Hartley 1997 p141).

Continued over

Group Variables	Explanation	Notes
Social, political or power status of the group. (Adler 2000; Dillon 1993)	The positioning of the group relative to other groups, or as a sub-group within other groups affects the group's and group members' power and behaviour. Group norms may overlap, dominate or conflict with other group norms, affecting member and group behaviour.	To illustrate how one set of norms may dominate over another, a team leader may dominate a reviewer according to hierarchical norms in a collaborative writing (discourse community) group. However, this expected behaviour may conflict with organizational norms if the team leader is a software engineer and the reviewer is the Managing Director. Levels of overlapping loyalties may exist (Hartley 1997 p21).
Group evolution	Groups are dynamic and change through different phases with time. Identifiable phases in group development are defined by group member behaviour, and may in turn affect group member behaviour.	Tuckman synthesizes research to describe five phases in group development: forming, storming, norming, performing and adjourning (Tuckman 1965, 1977 cited in Cole and Nast-Cole 1992). The first three phases appear relevant to group formation and the fourth to task completion (Hartley 1997 p54). Bales, using his IPA noticed that groups typically switched between task and maintenance (social and emotional) activities (Bales 1951 p131; also cited in Hartley 1997). Beck observed the need for group communications related both to writing task and group maintenance (Beck and Bellotti 1993; also cited in Mitchell 1996 p13).
Interdependencies (Sharples 1993 p58)	Dependence of one task on the completion of another. The dependence on each other's skills.	Argyle describes dependency as a key motivational drive in interpersonal behaviour with the central goal of obtaining and keeping nurturant and supportive relationships. He relates dependency to survival origins in childhood and to submissiveness as opposed to dominance (Argyle 1994 p9).

Appendix C: Individual Variables

Individual Variables	Description/Explanation	Notes
Knowledge (Hayes and Flower 1980)	of writing as an individual or in a group (Diaper 1993; Ford et al. 2000); of discourse community (Beaufort 2000); of intended audience; (Nystrand 1989); of discipline or area of expertise (Beaufort 2000; Ford et al. 2000; Odell 1985); of supporting technology (Ford et al. 2000); of medium of communication.	
Procedural and processing practices specific to writing task (Faigley 1985; Hayes and Flower 1980; Nystrand 1989; Odell 1985; Sharples and Pemberton 1988)	Different writers adopt different strategies or ways of organizing activities in the writing process, for example how they respond to text just created. (An individual's analytical skills must play a role in these processes.)	Flower and Hayes describe the process as recursive with main generating, organizing, translating and reviewing phases (Flower and Hayes 1980, 1981a&b; Hayes and Flower 1980). They allow for individuality giving examples of different strategies as *depth first*, *get it down as you think of it then review*, *perfect first draft* and *breadth first* (Hayes and Flower 1980 p20). Sharples describes the process as cyclic: planning, engaging, reviewing and reflection and also allows for individual's different strategies (Sharples 1999a). Nystrand's testing for writer/reader reciprocity falls into these reviewing and evaluative cognitive processes, but extends the paradigm to involve the reader and meaning as a social construct (Nystrand 1989).

Continued over

Individual Variables	Description/Explanation	Notes
Social roles, human nature and interrelationships are influential in team writing (Beck 1993; Diaper 1993; Kendrick 1998; Mitchell 1996; Sharples 1999b). The following individual attributes are measurable indicators of human nature.		
Intelligence	The ability to comprehend; to understand and profit from experience. An individual's intelligence is assumed to influence the individual's performance in any area, together with the individual's aptitude for the area.	Listed as one of the most commonly used personality traits by which people are judged by others (Argyle 1994 p 81).
Level of sociability (extraversion)	Level of sociability can be measured on a scale, and is partially (50%) accounted for by genes.	Listed as one of the most commonly used personality traits by which people are judged by others (Argyle 1994 p 81). Extraverts are sociable, positive, happy people, with strong social skills (Argyle 1994 p122). Extraversion is fairly stable with time and predictable over long time periods (Argyle 1994 p124). Argyle writes that in work-groups, "a pattern of informal social life develops as well" leading to group cohesiveness (Argyle 1994 p175).
Likeability	Agreeableness.	Listed as one of the most commonly used personality traits by which people are judged by others (Argyle 1994 p81). Liking has been measured by Moreno (cited in Hartley 1997 p128). His sociometry scale of spontaneous choice and sociometric preference is reported in Miller (Miller 1991 p376). The scale identifies friendship networks and is based on judgements on liking of other members and ranking according to preferences to work with them. Argyle writes that in work-groups, "a pattern of informal social life develops as well" leading to group cohesiveness (Argyle 1994 p 175).

Continued over

Individual Variables	Description/Explanation	Notes
Emotional stability		Listed as one of the most commonly used personality traits by which people are judged by others (Argyle 1994 p81).
Social competence	Social skills, empathy, social intelligence and problem solving, application of feedback to adjust behaviour ("translation"), assertiveness.	Assertiveness is listed as one of the most commonly used personality traits by which people are judged by others and also a factor contributing towards social competence (Argyle 1994 p 81).
Organizational role	Hierarchical position in an organization.	This can affect behaviour, e.g. equal status people not in the same group are often friendly and relaxed (Argyle 1994 p183). Argyle believes more can be predicted of the interpersonal behaviour between two people from their organizational levels than their personality traits (1994 p181-2).
Functional role	Combination of tasks the individual performs for the organization e.g. as a programmer, writer, administrator, supervisor etc.	
Socialisation in discourse community, functional and/or organizational group (Beaufort 2000; Faigley 1985; Ford et al. 2000; Odell 1985)	How "established" a member is or feels he or she is relative to other group members and ability to perform the required tasks.	Self perception as a group member (Tajfel and Fraser 1978 cited in Hartley 1997 p22). Whether worked with other team members before may accelerate socialisation.

Continued over

Individual Variables	Description/Explanation	Notes
Aspirations/individual goals or motivations (Argyle 1994; Ford et al. 2000)	Need for affiliation and need for dominance. Motivation related to task.	The need for friendship at work or the need to dominate over other people. Women have a greater need of affiliation at work than men (Argyle 1994 p10) and also self-disclose more often (p134). Argyle writes that in work-groups, "a pattern of informal social life develops as well" leading to group cohesiveness (Argyle 1994 p175). The balance of affiliation/dominance affects relationships. Someone preferring to be in a warm and superior relation will interact well with someone preferring to be in a warm submissive relation to others (Argyle 1994 p114).
Culture and/or ideology (Faigley 1985; Odell 1985)	Philosophies, beliefs, principles, traditions, etc. held by a larger group to which the individual belongs.	May include social class e.g. in Britain, or some religions, e.g. Hinduism. "Their whole way of life – their language, ways of perceiving, categorizing and thinking about the world, forms of non-verbal communication and social interaction, rules and conventions about behaviour, moral values and ideals, technology and material culture, art science, literature and history" (Argyle 1994 p184).
Gender and age (Ford et al. 2000)		

Appendix D: Social Variables

Social Variables	Description/Explanation
Infrastructure of production defining the discourse community or locales (Faigley 1985; Greenberg et al. 1999; Mitchell 1996; Nystrand 1989; Odell 1985)	Discipline or area of expertise; media of communication (FtF or networked teams, or combination of both) (Diaper 1993); available technology to support task and group maintenance activities (Diaper 1993; Sharples 1993).
Task specification	Document specification, type (instructive or persuasive, commercial or academic), genre, purpose (Dillon 1993), content, intended audience (Faigley 1985; Nystrand 1989), length, organization, deadline for completion (Beck 1993; Dillon 1993), importance of this task relative to other tasks in terms of organizational goals (Beaufort 2000).
Individual or Group characteristics and behaviour	See individual and group characteristics.
Organizational norms (Adler 2000)	Existing behaviour adopted and accepted by members of the organization. Levels of overlapping loyalties (Hartley 1997 p22).
Functional norms (Adler 2000)	Existing behaviour adopted and accepted amongst members with common functional roles within an organization.
Relative positioning of different group norms (organizational, functional, discourse community etc)	How one set of norms may dominate over another: for example, a team leader may dominate a reviewer following collaborative writing (discourse community) group norms. However, this expected behaviour may conflict with organizational norms if the team leader is a software engineer and the reviewer is the Managing Director.
Influence of other work on this work (Dillon 1993)	How other work impacts the group's ability to continue with this work, i.e. time management, dependency issues, such as on product development.
Previous texts (Faigley 1985; Odell 1985)	Existing documents or draft documents or instantiated text in the current document relating to the group's goal.
Organizational goals	Goals common to the organization and all its members, e.g. profitability.

Appendix E: Performance Measures

Item for performance measure	Description	Performance measure
Document	End product of the collaborative writing process.	Usability of document (different methods reported in Preston 2004). Readability measures (Faigley 1985) Flesch etc. to be used with caution (Hartley 2004). Accurate audience interpretation of intended meaning (Nystrand 1989). Appropriateness for audience (Faigley 1985; Sless, 2004). Appropriateness for audience context (Sless 2004). Topic progression and flow (Faigley 1985).
Process	The sequence of activities performed by the group members to achieve the group goal, i.e. to create the document.	Performance within timescale. Performance within budgeted resources. Success measured against individual, group and organizational goals. Success measured against equivalent groups. Perceived satisfaction of individuals, group, and organization (Aytes measured perceived process satisfaction (Aytes et al. 2002).

Appendix F: Mapping variables in the process and identifying measures

	Sub-variables	Sub-sub variables	Data type	Individual level data	Group level data	Data source
Input - Group Variables						
Group goal	Document type		Categorical		Category	Interview with project responsible
	Budget		Ratio		Currency unit	Interview with project responsible
	Time-scale		Ratio		Working days	Interview with project responsible
Group size			Ratio		No. in team	Interview with project responsible & e-mail data (below)
Group age			Ratio		Working days	Interview with project responsible
Group stability	Preconceived		Interval & descriptive		Score (e.g. 1=low; 5=high)	Interview with project responsible: stability & reasons
	Historical		Ratio		% membership change	Questionnaire & e-mail content analysis (see below)
Inividual 1-n	Knowledge/Exp.	of writing	Ratio	Years	Total & Av. & Variation	Questionnaire
		of team work	Interval	Score	Average & Variation	Questionnaire (Likert scale)
		of audience	Interval	Score	Average & Variation	Questionnaire (Likert scale)
		of subject matter	Ratio	Years	Total & Av. & Variation	Questionnaire
		of technology	Interval	Score	Average & Variation	Questionnaire (Likert scale)
		of comm. medium	Interval	Score	Average & Variation	Questionnaire (Likert scale)
	Qualifications	in writing	Interval	Category	Category frequencies (%)	Questionnaire (open question)
		in subject matter	Interval	Category	Category frequencies (%)	Questionnaire (open question)
	Cognitive strategies		Categorical	Category	Category frequencies (%)	Questionnaire (discovery/planner)
	Intelligence		Ratio	Score	Total & Av. & Variation	Questionnaire (Eysenck's IQ test)
	Extraversion		Ratio	% Score	Total & Av. & Variation	Questionnaire (Eysenck's personality test)
	Likeability		Ratio	% likes/dislikes	Total & Av. & Variation	Questionnaire (Moreno's likeability score 1934, cited in Miller 1991)
	Emotional stability		Ratio	% Score	Total & Av. & Variation	Questionnaire (Eysenck's personality test)
	Organizational role		Categorical	Category	Category frequencies (%)	Questionnaire
	Functional role		Categorical	Category	Category frequencies (%)	Questionnaire

Continued over

	Sub-variables	Sub-sub variables	Data type	Individual level data	Group level data	Data source
	Socialisation in discourse community		Interval	Score	Average & Variation	Questionnaire (Likert scale)
	Aspirations	Personal needs (to affiliate etc.)	Interval	Score	Average & Variation	Questionnaire (Likert scale of importance)
		Professional success	Interval	Score	Average & Variation	Questionnaire (Likert scale of importance)
		Financial reward	Interval	Score	Average & Variation	Questionnaire (Likert scale of importance)
	Culture/ideology		Categorical	Category	Category frequencies (%)	Questionnaire (ethnic origin)
	Gender		Categorical	Category	Category frequencies (%)	Questionnaire
	Age		Ratio	Years	Average & Variation	Questionnaire
Group role composition			Categorical	Category	Category frequencies (%)	Questionnaire
Group cohesiveness			Interval	Score	Average & Variation	Questionnaire Seashore's index (1954 cited in Miller 1991 p375)
Subgroup existence	Perceived		Ratio & Category		No. & type	Questionnaire
Group norms	Type		Descriptive		Categories	Interview & Questionnaire
	Level of influence		Interval	Score	Average & Variation	Questionnaire (Likert scale)
Group development						E-mail data with time (see below)
Group status			Descriptive			Interview with project responsible
Interdependencies			Descriptive		Categories	Questionnaire
Input - Social Variables						
Infrastructure of production			Categorical (technology type; geographically dispersed etc.)			Interview with project responsible
Task specification (type & how well specified)			Categorical & Interval			Interview with project responsible & Questionnaire
Organization (hierarchy of functional roles)			Descriptive			Interview with project responsible
Organizational norms			Descriptive			Interview & Questionnaire
Functional norms			Descriptive			Interview & Questionnaire
Relative positioning of norms			Ordinal			Interview & Questionnaire
Influence of other goals on this work			Descriptive			Questionnaire
Previous texts			Descriptive			Questionnaire
Organizational goals			Descriptive			Interview with project responsible

Continued over

	Sub-variables	Sub-sub variables	Data type	Individual level data	Group level data	Data source
Process Measures						
Project development			Descriptive	Events & date		Journal of key dates & e-mail analysis (see below)
E-mail Dynamics						
Group size			Ratio		No. in team	No. e-mail originators
E-mails per individual			Ratio	% frequencies		Sent e-mails/total e-mails
Exchange rate			Ratio	No. per day	Av. No. per day	Sent e-mails/time
Response times			Ratio	Av. response time	Av. response time	Time between receipt and response
Relative activities by functional role			Ratio	% activity	% activity	Role coding, calibrations & e-mail frequency
Relative activities by task			Ratio	% activity	% activity	Task coding, calibrations & e-mail frequency
Relative activities: task/group maintenance			Ratio	Ratio activities	Ratio activities	Task coding (van der Meij et al. 2004)
Evolution of role activity with time			Ratio	% activity v. time	% activity v. time	Role coding, calibrations & e-mail frequency
Evolution of task activity with time			Ratio	% tasks v. time	% tasks v. time	Task coding, calibrations & e-mail frequency
Directional flow in organization			Ratio	% up, down, same	Av. % up, down, same	Organization, e-mail direction coding & frequencies
Distribution lists			Categorical	% frequencies	% frequencies	No. of e-mails with same circulation lists
E-mail Content						
Key events & dates			Descriptive		Key event & date	Extracted from e-mail content
Group stability			Ratio		% membership change	Replacement, increase, decrease from e-mail content
E-mail size			Ratio	Av. no. words/segs.	Av. no. words/segs.	Counts per e-mail
Greetings content			Ratio	% freq. epistolary	Av. % freq. epistolary	% presence of social building units in e-mails
Formality			Interval	Score (0-2)	Av. Score (0-2)	Scores for formality per e-mail (Te'eni et al. 2001)
Use of affect			Ratio	% affect segments	Av. % affect segments	% segments expressing affect (van der Meij et al. 2002)
Question response rate			Ratio	% response rate	Av. % response rate	Question/reaction coding (van der Meij et al. 2002)
Communication threads			Ratio	Threads/e-mails	Av. no. threads/e-mails	Reaction segment coding (van der Meij et al. 2002)

Continued over

	Sub-variables	Sub-sub variables	Data type	Individual level data	Group level data	Data source
Output measures						
Document usability	Group goals		Interval	Scores		Questionnaire (usability issues)
	User feedback		Interval	Scores		Questionnaire (usability issues)
Productivity measures	Individual goals		Interval scale	Scores		Questionnaire (Likert scale)
	Group goals		Interval	Scores		Questionnaire (Likert scale)
	Org. goals	budget	Interval	Scores		Questionnaire (Likert scale)
		time-scale	Interval	Scores		Questionnaire (Likert scale)
		cf. other groups	Interval	Scores		Questionnaire (Likert scale)

Appendix G: Mapping variables for interpreting social and task dimensions

Task: Variables likely to tell us something about the professional (organizational and functional) aspects of the process

Variables	Sub-variables	Sub-sub variables	Data type	Individual level data	Group level data	Data source
Individuals' Professional Level	Knowledge/Exp.	of writing	Ratio	Years	Total & Av. & Variation	Questionnaire
		of subject matter	Ratio	Years	Total & Av. & Variation	Questionnaire
	Qualifications	in writing	Interval	Category	Category frequencies (%)	Questionnaire (open question)
		in subject matter	Interval	Category	Category frequencies (%)	Questionnaire (open question)
	Organizational role		Categorical	Category	Category frequencies (%)	Questionnaire
	Functional role		Categorical	Category	Category frequencies (%)	Questionnaire
	Socialisation in discourse community		Interval	Score	Average & Variation	Questionnaire (Likert scale)
	Socialisation in organization		Interval	Score	Average & Variation	Questionnaire (Likert scale)
	Socialisation in functional community		Interval	Score	Average & Variation	Questionnaire (Likert scale)
Organization (hierarchy of functional roles)			Descriptive			Interview with project responsible
Directional flow in organization			Ratio	% up, down, same	Av. % up, down, same	Organization, e-mail direction coding & frequencies
Organizational norms			Descriptive			Post project questionnaire/interview
Functional norms			Descriptive			Post project questionnaire/interview
Project development			Descriptive	Events & date		Journal of key dates & e-mail content analysis
Relative activities by functional role			Ratio	% activity	% activity	Role coding, calibrations & e-mail frequency
Relative activities by task			Ratio	% activity	% activity	Task coding, calibrations & e-mail frequency
Evolution of functional role activity with time			Ratio	% activity v. time	% activity v. time	Role coding, calibrations & e-mail frequency
Evolution of task activity with time			Ratio	% tasks v. time	% tasks v. time	Task coding, calibrations & e-mail frequency

Continued over

Social: Variables likely to tell us something about the group maintenance (social or non-task oriented) aspects of the process

Variables	Sub-variables	Sub-sub variables	Data type	Individual level data	Group level data	Data source
Individuals' Group Integration	Likeability		Ratio	% likes/dislikes	Total & Av. & Variation	Questionnaire (Moreno's likeability score, 1934 cited in Miller 1991 p 376)
	Socialisation in group		Interval	Score	Average & Variation	Questionnaire (Likert scale)
	Group role		Categorical	Category	Category frequencies (%)	Questionnaire
Group cohesiveness			Interval	Score	Average & Variation	Questionnaire Seashore's index (1954 cited in Miller 1991 p375)
Subgroup existence			Ratio & Category		No. & type	Questionnaire
Distribution lists			Categorical	% frequencies	% frequencies	No. of e-mails with same circulation lists
Question response rate			Ratio	% response rate	Av. % response rate	Email question/reaction coding (van der Meij 2004)
Communication threads			Ratio	Threads/e-mails	Av. no. threads/e-mails	Email reaction segment coding (van der Meij 2004)
Exchange rate			Ratio	No. per day	Av. No. per day	Sent e-mails/time
Response times			Ratio	Av. response time	Av. response time	Time between email receipt and response
Greetings content			Ratio	% freq. epistolary	Av. % freq. epistolary	% presence of social building units in e-mails
Formality			Interval	Score (0-2)	Av. Score (0-2)	Scores/e-mail (Eggins & Martin 1997; Te'eni et al. 2001)
Use of affect			Ratio	% affect segments	Av. % affect segments	% email segments expressing affect (van der Meij 2004)
Group status			Descriptive			Interview with project responsible
Group norms			Descriptive			Post project questionnaire/interview
Relative positioning of norms			Ordinal			Post project questionnaire/interview
Group stability			Ratio		% membership change	Replacement, increase, decrease from e-mail content
Relative activities: task/group maintenance			Ratio	Ratio activities	Ratio activities	Email task coding (van der Meij 2004)